THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR ....Alexander Leonard HILLYARD........

TITLE OF THESIS ....Stimulus Complexity During Original

........Learning and Generalization........

DEGREE FOR WHICH THESIS WAS PRESENTED ....Doctor of Philosophy....

YEAR THIS DEGREE GRANTED ....Fall, 1979........

DATED ....Oct 3.......1979

THE UNIVERSITY OF ALBERTA

STIMULUS COMPLEXITY DURING ORIGINAL LEARNING AND GENERALIZATION

BY

© ALEXANDER LEONARD HILLYARD

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

IN

EDUCATIONAL PSYCHOLOGY

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1979

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled STIMULUS COMPLEXITY DURING ORIGINAL LEARNING AND GENERALIZATION submitted by ALEXANDER LEONARD HILLYARD in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Educational Psychology.

ABSTRACT

The present study had three main purposes; first to compare the effectiveness of two dimensions of instructional programming, diversity and number of concept exemplars, when teaching relational concepts to mentally retarded preschool children; second, comparing the effectiveness of these dimensions in facilitating generalization following a criterion for mastery; and third to examine the relation between acquisition and generalization for minimal instructional costs and maximum generalization.

The two dimensions under investigation were diversity of concept exemplars and number of exemplars from a concept class. Diversity of exemplars had two levels (IER); level one had members of a concept pair varying on the relevant dimension distinguishing the concept and on one irrelevant dimension (position); level two had members varying on the relevant dimension defining the concept and on three irrelevant dimensions (position, color, form). All other irrelevant dimensions between concept member pairs at each level remained constant. Number exemplars (NE) also had two levels; level one had two different stimulus pairs during training, while level two had four different stimulus pairs during training.

Four experimental conditions were combined from these dimensions; condition one, level one IER and level one NE: condition two, level one IER and level two NE; condition

three, level two IER and level one NE; condition four, level two IER and level two NE. Experimental conditions were combined into two groups (conditions one and two, conditions three and four) and were trained in two sets of blocked trials. The order of introduction for condition groups and daily training blocks were counterbalanced across subjects.

A pilot study with two moderately retarded subjects revealed a flaw in the instructional procedure which created a response set. This response set was demonstrated in an A-B-A design. The instructional procedure was subsequently changed for the main experiment.

Experimental subjects included three preschool aged, moderately retarded children (two boys, one girl) and one preschool age severely retarded boy. Each subject was taught four sets of polar opposite concepts, one set in each experimental condition. Concept pairs were presented in a simultaneous format.

The results indicated, with the exception of one subject in one experimental condition, that all subjects acquired all four concept pairs in the four experimental conditions. Number of exemplars was shown to produce greater trials to criterion during acquisition as it increased from level one to level two, however, there was no major difference in generalization. Diversity of exemplars also produced greater trials to criterion as it increased from level one to level two, however level two produced greater generalization. The

results suggest that diversity of exemplars rather than number of exemplars is the crucial dimension in programming for generalization.

A comparison of instructional costs, in terms of trials to criterion, between acquisition and generalization suggests that fewer exemplars with high diversity results in fewer trials to criterion and similar amounts of generalization when compared to more exemplars with high diversity.

Plausible reasons for the results obtained were discussed and their implications for future research in the area of instructional programming for concept training and generalization were also suggested.

ACKNOWLEDGEMENTS

The author wishes to express sincere appreciation and gratitude to the following individuals for their assistance during the preparation and completion of this thesis:

Dr. Gerry Kysela, my thesis advisor, for his extensive advice and guidance; Dr. David Baine, Dr. Robert Short, Dr. Steven Carey and Dr. Gary Holdgrapher for their constructive criticisms; Dr. Anne Rogers-Warren, my external examiner, for her thorough analysis of the thesis and enlightening discussion of the topic area.

I would also like to express my sincere appreciation to the following colleagues and friends who both knowingly and unknowingly helped me through a family difficulty during the final preparation stages and without their support this thesis may not have become a reality.

By name they are James and Jennie Chapman, Linda and Stew McDonald, Julie and Bruce Taylor, Janie and Greg Lysack, Jody and Ian Thumlert, Gerry Kysela and David Baine.

Special recognition as words fail me, is expressed to my mother and father and grandmother Marjorie Jean Brodie Smith to whom this thesis is endearingly dedicated.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The present investigation was concerned with the effects of stimulus complexity during original learning and resultant generalization for concept teaching with preschool mentally retarded children. When designing concept teaching programs for this population it is important to describe the instructional parameters or task variables which influence generalization and retention of concept usage.

Becker, Engleman and Thomas (1975) employing a direct model of instruction for concept teaching have demonstrated the effectiveness of their approach when applied to disadvantaged preschool age children. Concept teaching through direct instruction involves the two processes of teaching discrimination and generalization (Becker, Engleman and Thomas, 1975). A child learns to discriminate relevant stimulus dimensions defining a concept from irrelevant stimulus dimensions not representing the concept. Secondly, generalization involves identification of novel instances of the concept which were not present during teaching.

There is a lack of research describing how this approach to concept teaching can be applied to preschool mentally retarded children. It would appear an important applied research

question to evaluate the components of a direct model of concept teaching with a population for whom it has potential usefulness.

There are a number of task variables within a direct teaching approach which can have potential influence on a learner's performance. Two of these task variables are the nature and relation of relevant and irrelevant cue dimensions when arranging stimulus items during teaching. These stimulus characteristics are concerned with the amount and nature of information provided to a student when presenting instructional programs.

The research on concept formation has focussed on normal children and adults in relation to the role of stimulus complexity during original learning; little study has been done relating to the generalization of concept usage. Stokes and Baer (1977) in a review of the instructional or program conditions which influence generalization suggested two parameters which have considerable relevance to the area of concept teaching. The first parameter is concerned with the diversity of exemplars presented during original learning and corresponds to the previously described notion of stimulus complexity involved in the variation of the number of relevant and irrelevant dimensions present. The second parameter is concerned with the number of exemplars presented during original learning.

This thesis has attempted to address the above problem by demonstrating the influence of diversity and number of exem-

plars on initial learning and generalization of concept usage with preschool mentally retarded children. Specifically this was done by indicating the independent and interactive effects of these two program dimensions on initial acquisition, generalization and retention of concept usage. The results have implications for the design of further research studies in concept teaching and general program considerations when designing instructional concept teaching programs for this population.

CHAPTER II

LITERATURE REVIEW

Introduction

The terms concept and conceptual behavior can have a wide variety of meanings. Flavell (1970) attempting to find a common definition of the term concept by reviewing existing definitions concluded that the only common feature among definitions was their differences. It appeared that no one definition could satisfy the constraints of the other. Gagne (1966) attempting to arrive at a common definition of concepts disagreed with Flavell's (1970) conclusion. Gagne (1966) suggested "despite differences in the language used to describe a concept there is considerable agreement among research psychologist's as to what these words mean" (Klausmeier and Harris, 1966, p. 82). He isolated the following three general properties common to all definitions:

1) a concept is an inferred mental process;

2) learning of a concept requires discrimination of stimulus objects (distinguishing positive and negative instances;

3) performance which shows that a concept has been learned consists of the learner being able to place an object in a class.

There are three general theoretical approaches to the

study and explanation of concept learning. One originates within the stimulus-response associationist approach to basic learning processes. A second is the more cognitive approach emphasising the role of learner initiated hypothesis and strategies. The third is presented by proponents of information processing approaches to human learning. In each of these approaches there is not one theorist or theory representing an approach but rather a constellation of approaches following similar general guidelines or underlying principles. Each of these approaches are described briefly.

Association Theory:

Association theories of concept learning are either referred to as S-R association theories or S-R mediational theories. These two theories are similar enough to be classified within one approach, however, there is one important distinguishing factor.

Concept learning tasks are viewed as a series of instances or exemplars containing attributes which are relevant or irrelevant to identification of the concept. S-R theorists suggest that an association, between the relevant attributes defining the concept and a response of a subject observing these relevant attributes is formed as a result of environmental feedback. The association between responses to relevant attributes are strengthened through repeated pairing and environmental feedback and reinforcement whereas the association between responses to irrelevant attributes

is weakened through repeated pairing without environmental feedback and reinforcement. Concept learning is viewed as a special case of discrimination learning and stimulus generalization through the process of differential reinforcement and environmental feedback.

An individual has learned a concept when he can identify new examples of the concept class. The theory suggests that what is learned in concept learning is an association between similar stimuli and the same response. This association is learned through differential reinforcement, and nothing else is suggested to account for concept learning.

Additional S-R mediational thories have been developed to account for concept learning. These approaches suggest an association first between a common covert stimulus and secondly between the covert stimulus and an overt response. The defining characteristic then of mediation theory is an internal representation of an external stimulus which acts as a cue for further overt behavior.

The mediational component within S-R association theories was postulated due to the expressed inadequacies of S-R association theory to account for some classes of conceptual behavior. These classes of conceptual behavior were those that have no observable common elements or attributes in the examples forming the concept class. A common example of such a concept class is food. Instances of food do not have common physical features from which one could associate. The common response to instances of food is eating and it is this

response which acts as a mediator to identify examples as members of the concept class. S-R mediational theories are seen as more complex than S-R association theories in addition to being more flexible in accounting for the learning of a broader range of concepts.

Hypothesis Testing Theories:

The important distinction between hypothesis testing theories and S-R theories is the role attributed to the learner. S-R theories suggest that the learner's responses are under the direct control of external stimulus events or internal mediating cues. Hypothesis-testing theory on the other hand suggests that the learner is actively engaged in the learning process through formulating hypotheses about external stimuli and actively testing them. External feedback to the learner results in confirmation, deletion or addition to his existing hypotheses.

The most cited examples of this approach to studying concept learning is the work of Jerome Bruner (Bruner et al. 1956). Bruner's work was primarily concerned with rule learning and the types of strategies subjects employed in solving rule based conceptual problems. It is difficult to contrast S-R approaches which do not deal with rule learning directly and hypothesis testing which is primarily concerned with rule learning or rule identification. A further discussion of S-R theories and hypothesis testing theories will be developed in a later section outlining the model of concept learning used in the present research.

Information Processing Theories:

Information processing theories make an analogy between man and computers. The analogy is based on the fact that both receive external input, produce an observable response and the resultant response is attributed to internal workings or processes. These theories are an extension of hypothesis testing approaches in the sense that they place emphasis on internal cognitive activities.

Computer programs are constructed to simulate what is inferred to occur in the human brain during the concept learning. The general approach makes several assumptions about human mental processes which do not seem to be supported in experimental literature. These assumptions are perfect memory processes of storage and retrieval and error-free processing strategies when attacking conceptual problems.

These three general approaches have been used to describe and study how concepts and concept learning occurs. As Horton and Turnage (1976) point out many theories within each approach have been developed within the context of one type of conceptual problem (i.e. attribute identification versus rule learning) and may not be directly interpretable as an explanation of other conceptual problems. In addition these theories have been developed to account for specified concept classes and not for abstract concepts or rule learning (i.e. S-R association theories). This does not mean that one theory or approach cannot account for other forms of conceptual problem learning but rather

that they have not studied them directly.

Taken in a continuum the three approaches can be viewed as moving from a simple to more complex explanation of concept learning. However, each approach has studied different levels of concept and rule learning. It is possible to suggest as Horton and Turnage (1976) imply, that each approach explains different levels of concept learning and that each approach exists as an adequate explanation in relation to the concept class under study. An alternate possibility is that more complex explanations which move further into inferred unobservable phenomena only take us further away from a more parsimonious explanation. The present research assumes a S-R association or operant approach to concept learning. It is suggested that such an approach can account for all classes of conceptual behavior through the established principles of discrimination, stimulus generalization and differential reinforcement. The rationale for adopting this approach is discussed in the next section which describes the assumed model of concept learning.

Assumed Model of Concept Learning

An operant approach to concept learning describes concepts as discriminated operants controlled by a class of discriminitive stimuli. The controlling variable is differential reinforcement for responding or failing to respond in the presence or absence of discriminitive stimuli.

Keller and Schoenfeld (1950) describe the operant approach and emphasize observable behavior;

> What is a concept? This is another term which has come into psychology from popular speech, carrying with it many different connotations. We shall have to be careful in using it, remembering that it is only a name for a kind of behavior. Strictly speaking one does not have a concept, just as one does not have extinction, rather, one demonstrates conceptual behavior by acting in a certain way. Our analysis should really start with a different question: What type of behavior is it that we call conceptual?
>
> (Keller and Schoenfeld, 1950, p. 154)

As discussed earlier, concepts and concept formation can mean different things to different researchers. One way of describing an operant approach is to contrast it with other approaches. Opposing approaches have employed either a cognitive model (Hunt 1962, Osgood 1953) or developmental model (Kendler and Kendler 1959) to study and explain concept formation. Operant approaches differ from both these models on their definition of conceptual behavior, explanation of controlling variables and developmental changes in concept formation.

<u>Definition of Conceptual Behavior</u>

Definitions within an operant approach do not make reference to assumed or inferred mental processes but rather refer to "conceptual behavior" in terms of stimulus and response dimensions. The experimental definitions of Goldiamond (1962); Kendler and Kendler (1959); Kendler (1961); Osgood (1953); Spradlin, Cotter and Braxby (1973) agree that conceptual behavior is demonstrated through a similar response to a class of dissimilar stimuli.

Spradlin and Dixion (1976) felt that these earlier

definitions were too restrictive and further developed the definition by suggesting that it is the relationship between and interchangeability of stimuli which determines membership in a concept class. They suggest that "a concept consists of stimuli or events which are substitutable for each other within a given context. We assume if two or more stimuli are established as functionally equivalent (or substitutable) through reinforcement and training in one condition, there is an increased probability that they will be functionally equivalent in a second, even without direct reinforcement or training in that situtation" (Spradlin and Dixion, 1976, p. 555-556). Such a definition based on association of responses through reinforcement establishes response equivalents which can account for concept classes with no observable distinguishing attribute but rather a common response. This definition allows for increased flexibility when accounting for the establishment of different concepts within an operant approach.

The two most vocal critics of this definitional approach are Osgood (1953) and Hunt (1962). Osgood (1953) believes that concept learning requires an abstraction process which is unique to humans. He never really defines what this abstraction process is or how it occurs but suggests that if conceptual behavior is based on discrimination and generalization alone then animals can learn concepts. Osgood (1953) referred to Field's (1932) study on the concept of triangularity in the white rat, which he suggests was an example

of complex discrimination learning and not concept formation. Field (1932) trained a rat to discriminate between two doors when jumping from the Lashley jumping stand. Employing a negative reinforcement procedure the door which avoided an aversive situation was marked with a triangle. During the training procedures the triangle was varied with new dimensions of size, color, position and shading.

The rat's similar response to a series of dissimilar stimuli would appear to meet Osgood's (1956) definition where a concept is a set of stimuli which controls the same response. Osgood (1953) suggested that the rat's behavior did not denote conceptual behavior as it may not be responding to the abstract concept of triangularity. This suggestion by Osgood centers on the question of what did the rat learn? Within the behavioral limits of a similar response to dissimilar stimuli the rat did exhibit conceptual behavior however, it is impossible to decide if this behavior also reflects abstraction of the concept triangle.

According to Osgood's definition, abstraction is defined by response characteristics and as a result Field's rat would appear to have abstracted the concept triangle. However, if abstraction is argued to be an internal unobservable process within Osgood's (1956) definition, then it can be argued that Field's rat did not abstract the concept. The attention of any definition to assumed internal unobservable mental processes suggests that any subject population can never really show conceptual behavior as one

cannot directly observe mental processes. If on the other hand one infers mental processes from behavior then conceptual behavior does reflect abstraction of a concept. This applies equally to Field's rat as it does to human subjects.

Osgood (1953) also suggested that Field's rat may not respond in a similar way to new examples of a triangle (i.e. three people forming a triangle, three cornered block etc). This point brings up the question of when a concept is formed. Miller (1976) points out that during concept learning a subject may demonstrate conceptual behavior beyond the stimulus items during training but not to all possible items representing the concept class.

Human and infrahuman subjects demonstrating conceptual behavior at different levels of stimulus complexity may be a result of original learning conditions and breadth of parameters defining a concept class. If Field's rat displayed similar responses to stimulus changes in the concept class as suggested by Osgood (1953) when would the criteria for abstraction be satisfied? Field's (1953) rat was responding to dissimilar stimuli and therefore Osgood's (1953) question is almost impossible to answer experimentally unless it is further operationally defined. All that can be argued is the observable data showing controlling external stimuli resulting in the behavior of a rat which is similar by definition to conceptual behavior in humans.

A second point of criticism arises from Hunt's (1962)

definition of a concept and discussion of an operant approach. Hunt's (1962) definition differs from that of an operant approach on the response characteristics which define conceptual behavior. His definition suggests that conceptual behavior is contingent upon displaying a verbal statement describing the relationship between defining stimuli. As a result Hunt's definition excludes infrahuman subjects and nonverbal human subjects suggesting conceptual behavior is unique to humans with specific communication skills.

The importance placed on verbalization of a rule in concept formation may be an unwarranted criteria. Verbalization of a rule may be another level of conceptual behavior demonstrating the use of a verbal attribute associated with a nonverbal response. Hull (1920, in Mussen 1965), and Piaget (1956 in Flavell 1972), report examples of experimental studies where subjects demonstrated conceptual behavior nonverbally (through manipulation of task stimuli) but did not or were unable to describe the rule. The two response modes verbal and nonverbal may be different skills which are related but reflect different levels of concept attainment.

Controlling Variables in Concept Learning

An explanation of existing approaches to concept formation and developmental changes in concept formation will be presented with an interpretation based on the controlling influence of differential reinforcement. An operant approach

defines differential reinforcement as the major controlling factor in achieving stimulus control. Differential reinforcement is also referred to as the major controlling factor when establishing more complex stimulus classes (i.e. concepts). Other variables influencing conceptual behavior are antecedent simulus dimensions such as stimulus presentation or stimulus sequencing. These variables have been studied primarily in the area of errorless discrimination learning (Terrace 1963).

Errorless discrimination procedures have been used to train complex discrimination in normal and retarded children (Sidman 1966, Sidman and Cresson, 1971). The present research is not employing an errorless procedure and this research approach will not be reviewed.

As described earlier mediational theories within an S-R approach and cognitive theories were developed to account for a suggested inadequacy in S-R association theory. This inadequacy was suggested on the grounds that S-R association theories could not account for concept formation when stimuli in the concept class do not have observable common defining dimensions. Concepts which do not share all the same dimensions or properties in common are referred to as disjunctive concepts.

Goldiamond (1962) describes an example of a disjunctive concept (stopping response) which illustrates how this concept class can be formed through differential reinforcement

and stimulus generalization. A person will display "stopping behavior" in the presence of a red light, stop sign etc., as a function of previous consequences to stopping or non stopping. The degree to which new members within the class (discriminitive stimuli (S^D) to stop) come to control the same response is a function of stimulus similarity to other stimuli in the class and consequences which follow it.

Spradlin et al (1976), studied the role of response equivalence between stimulus items when one stimulus was associated with an auditory cue. They found that an auditory cue associated with two members of a stimulus class came to elicit the same response for other members of the stimulus class. The stimulus items within the class were dissimilar on the dimensions of color and shape. Spradlin et al (1976) extended their results to suggest an explanation of the way in which receptive language items may be learned in childhood. The example used was a situation where a father gives his son a new toy to play with but does not tell him it is a toy. The son plays with the toy and is later requested by father to put all his toys away. The son picks up all his toys including the new toy and puts them away. The common response of playing, previously associated with the word toy is generalized to the new unlabelled "plaything" and it is included in the concept class toy.

It could be suggested that both examples from Goldiamond and Spradlin represent situations where a response mediates another response. An alternative plausible explanation is

put forth through stimulus generalization and association learning resulting from differential reinforcement.

Green (1955) and Namikas (1967), provide examples of research studies showing different acquisition rates for concept identification as a result of differential reinforcement schedules. Sherman et al (1967), working with young children initially established a concept class in a match to sample task through differential reinforcement then reversed responses to concept classes by reversing differential reinforcement. These studies illustrated how differential reinforcement controls the rate and direction when learning conceptual behavior.

Strategies used by subjects in concept learning situations are suggested to be beyond the framework of operant analysis and explanation as Duse and Hulse (1968) comment,

> it is a mistake to describe the learning of concepts as nothing more than a kind of passive process of discrimination, because we know that the behavior depends upon more than this, There is for example, the matter of hypotheses in concept learning. Even more to the point there is the matter of strategies. The use of strategies clearly lifts concept learning out of the domain of simple discrimination learning (p. 422).

Bruner et al (1956) studied and isolated different strategies (i.e. conservative or focusing) a subject may display when learning a concept. The different strategies were defined as an individual's stimulus choices when given feedback on previous choices. A subject's response characteristics were defined in relation to strategies rather than an explanation of previous learning history made up from differential consequences to stimulus items. An ob-

vious and just as plausible an explanation is that "Bruner's strategies" guiding response are learned response characteristics from previous stimulus-response learning based on differential reinforcement.

If a subject has a previous history where focusing on individual attributes has produced a greater probability of task completion over gambling strategies where two or more attributes are changed over each selection, then it seems reasonable to suggest he will choose a focusing strategy. The reverse is true for a gambling strategy depending on the previous learning history.

Other researchers, (Kendler and Kendler, 1959, 1967; Kendler, 1971) have suggested that an individual's strategy when learning concepts changes with age and the acquisition of language.

Kendler and Kendler (1959, 1967) generally found that children who learn initial discrimination problems easily and quickly learn the reversal problem but will have more problems with the nonreversal shift. Their results were interpreted as suggesting a mediation strategy employed by children to account for the differences in performance.

An alternative explanation could rest on the observed effects of partial reinforcement and extinction. During reversal shifts the new S^D or correct responses were not previously reinforced while the old S^D's, during original learning were reinforced on a continuous schedule. On the other hand, nonreversal shifts represent a situation of am-

biguity to the subject where during the shift, half of the previously continuously reinforced S^D's are now reinforced some of the time (partial schedule) and extinguished the remainder of the time.

The reversal shift represents a situation of extinction following continuous reinforcement of old responses and continuous reinforcement of new responses. Nonreversal shifts represent a situation where previous responses continuously reinforced are now only partially reinforced. The difference between these two response patterns to shift problems can be interpreted as a result of response persistence to different schedules of reinforcement and extinction. Partial reinforcement serves to increase response strength where extinction serves to weaken it (Skinner, 1953).

Summary

This section has attempted to show that concept learning explained within an operant approach is a viable alternative to opposing explanations of concept formation. As mentioned earlier, another controlling variable in an operant approach to concept learning is the presentation of antecedent events. This refers to the complexity of stimuli (amount of information presented), defining relationships between stimuli and the manner in which stimuli are presented.

The next section will outline research literature which has investigated these antecedent variables on original learning and resultant generalization. This research stems pri-

marily from the study of concept identification learning represented in the work of Lyle Bourne and others.

Literature Related to Antecedent Variables During Original Learning and Generalization

Bourne (1966) has outlined the task variables investigated within concept identification research and how they contribute to individual performance. Task variables are defined as "ecological conditions under which the subject must work to produce solutions to conceptual problems" p. 45. Task variables as a group are comprised of response variables, stimulus factors, information feedback and temporal factors. The task variables of interest for the present review and proposed research are stimulus factors. Stimulus factors generally refer to either the mode in which stimuli are presented (simultaneous or successive) or the complexity of stimuli presented in relation to the amount of information contained (e.g. amount of relevant and irrelevant information). More specifically the present research is concerned with the effects of stimulus complexity such as the amount of relevant and irrelevant information contained in instances of the concept class and number of instances from the concept class, as antecedent instructional parameters which influence original learning and generalization of concept usage. Literature investigating stimulus complexity will be discussed under two topical headings, one pertaining to the role of stimulus complexity during original learning while the second discusses the effects of stimulus complexity on generalization.

Variations of Relevant and Irrelevant Dimensions during Original Learning

There has been considerable research reported which investigated the independent and interactive affects of relevant and irrelevant dimensions during original learning of concept identification. Concept instances are represented by a set of dimensions which can have two or more values. Relevant dimensions are those stimulus cues which are crucial to concept identification. Other dimensions present in the stimulus cues which do not define the concept are referred to as irrelevant as their presence is not crucial for concept identification.

Stimulus variability within relevant and irrelevant dimensions are referred to as inter and intra dimensional variability. Interdimensional variability refers to stimulus complexity in terms of the number of different dimensions on which a concept instance can vary. This is defined operationally as the number of relevant and irrelevant dimensions. For example a concept instance may have one or more relevant dimensions and one or more irrelevant dimensions. Intra-dimensional variability refers to stimulus cue complexity where there are a number of values for each dimension. This form of variability may be assigned to either relevant or irrelevant stimulus dimensions. For example a concept may have size as a relevant dimension with two or more values (Size 1, Size 2, Size 3 etc) in addition to an irrelevant dimension (form) with two or more values (circle, square, triangle).

The effects of interdimensional variability for concept identification tasks has been well documented (Bourne and Haygood, 1959, 1961, 1961). Bourne and Haygood (1959, 1961) investigated the independent and interactive effects of relevant and irrelevant dimensions on performance in a concept identification task. They employed a receptive paradigm with successive presentation where multiple concepts were classified into one of four groups. Their experimental groups differed on the number of relevant and irrelevant dimensions present to define each concept. Their results showed consistently that subject performance improves as the number of relevant dimensions increases and the number of non-redundant irrelevant dimensions decreases. Bourne and Haygood (1961) replicated their earlier work with unidimensional and conjunctive concept problems while using redundant irrelevant stimulus cues rather than non redundant. They found similar results to their earlier research (1959, 1961) showing decreased performance as irrelevant dimensions increased. The decrease in performance was not as equivocal with non redundant irrelevant stimulus cues. The amount of irrelevant information has a negative effect on performance however, the form it takes will also dictate the extent of this negative effect. Their results also indicated that the increase in irrelevant dimensions produced a greater decrease in performance for the conjunctive over the unidimensional rule.

Battig and Bourne (1961) investigated the independent

and interactive effects of intra and inter dimensional stimulus variability on performance within concept identification tasks. Their study employed a receptive paradigm where subjects were informed of the conjunctive rule governing the concept (two relevant dimensions) in addition to descriptions of stimulus dimensions and possible values of each. Four concepts were presented on a screen in front of the subject where the task was to classify the stimulus object into one of four categories. A non-correction procedure was used in which immediately following a subject's response a light corresponding to the correct choice "lit up" one of four buttons corresponding to the correct response. A 3x3 factorial design represented all possible combinations of three conditions for interdimensional variability (1,2,3 irrelevant dimensions) and three conditions of intra dimensional variability (2,4,5 different stimulus values within each of the relevant and irrelevant dimensions). The four concepts were defined as two relevant dimensions (form and width) for the geometric shapes triangle and quadrangles. The dependent measures in the study were mean number of errors and trials to solution. Their results showed an increase in trials and errors to solution as each level of either intra or interdimensional variability increased. Both these sources of variability were significant at the .01 level. These results showed a similar significant ($p < .01$) linear trend as was observed for increases in irrelevant dimensions (Bourne

and Haygood, 1959, 1961). The results were more dramatic as the higher conditions interacted for both intra and inter dimensional variability. Battig and Bourne (1961) interpreted their results as fitting the Bourne and Restle (1959) concept identification model which suggests that the number of errors to solution is determined by the proportion of total cues relevant to problem solutions. An increasing intradimensional variability would be predicted to contribute to task complexity in a similar manner as interdimensional increases have shown.

Haygood, Harbert and Omlor (1970) investigated the role of intradimensional variability in concept identification tasks. They employed a receptive paradigm using a successive presentation with exemplars and nonexemplars. Independent variables were two levels of concept complexity (one relevant and two relevant dimensions) and three levels of intradimensional variability for irrelevant dimensions (e.g. three different-letter sets). The task consisted of choosing instances of a concept where stimulus material used was a set of five letters typed on a 3x5" card. In the one relevant dimension one position and letter were relevant while the two relevant dimension used two different letters and positions as relevant cues. The levels of intradimensional variability were two, four and six respectively. In two level problems each letter position could take on two possible letters, one of four in four level problems and one of six in

six level problems. There were an equal number of positive and negative instances for each condition and the criterion for task solution was sixteen correct responses. Two experiments were completed in which the results showed a decrease in trials to criterion as the number of levels in intradimensional stimulus variability increased. The two level relevant dimension appeared harder however, the difference from one level problems was not significant and there were no interactive effects.

When comparing their results to Battig and Bourne (1961) Haygood et al (1970) suggested that their results were not directly contradictory due to different manipulations. Battig and Bourne (1961) manipulated inter and intradimensional variability on both relevant and irrelevant dimensions while Haygood et al (1970) only manipulated intradimensional variability on irrelevant dimensions. Haygood et al (1970) suggested that Battig and Bourne's (1961) results may be limited to situations where attribute dimensions defining both categories of relevant and irrelevant dimensions increase in complexity rather than only one.

Haygood et al (1970) cautioned the interpretation of their results as they may be specific to the type of rule used for concept identification. They employed the affirmation and conjunctive rule which places considerable emphasis on the role of positive instances (Hovland 1952) and the possibility exists that rules which require attention to

negative instances for solution (conditional, joint denial) may show a decrease in performance as intradimensional variability increases for irrelevant dimensions.

Chumbley, Lau and Haile (1971) replicated the Haygood et al (1970) findings using a similar task format with conjunction concepts. It would appear that when the relative complexity of the task does not increase and irrelevant dimensions are interchanged through intradimensional variation an increase in performance results. Although Haygood et al (1970) and Chumbley et al (1971) did not specifically investigate subject selection strategies in concept attainment they both suggest their results are supportive of the wholistic positive focus strategy. This strategy suggests that a subject attends to the first positive instance presented and ignores all negative instances, using the positive instance as his working hypothesis. On subsequent trials the subject reduces his "whole" strategy by observing stimulus changes in the positive instances. This process reduces his hypothesis until only relevant dimensions are left (Haygood et al 1970).

Summary

The literature reviewed suggests that stimulus complexity does effect original learning, however, there remains some controversy over how different stimulus parameters contribute to this effect. The two stimulus parameters investigated were intra and inter stimulus variation on relevant and irrelevant

dimensions.

Generally it was found that as stimulus complexity increases for both relevant and irrelevant dimensions (Bourne and Haygood 1959, 1961) subject performance decreases, while increasing either intra or inter dimensional variability on relevant and irrelevant dimensions (Battig and Bourne 1961) also results in performance decrements. Alternatively, Haygood (1970) and Chumbley et al (1971) found that increasing stimulus complexity for irrelevant dimensions when the complexity of relevant dimensions remains constant results in fewer performance trials to solution.

The differences between these two general findings may be due to the type of conceptual rule underlying concept identification tasks in addition to the mode of stimulus presentation. Bourne and Haygood (1961) used disjunctive concepts which as Hovland (1952) points out requires attention to negative instances while Haygood et al (1970) and Chumbley et al (1971) used conjunctive and unidimensional concepts which do not place as much emphasis on negative information.

One could predict that stimulus complexity operationally defined by inter and intra stimulus variability will facilitate learning where the rule defining the concept does not place emphasis on negative instances for identification of positive instances. Increased complexity of irrelevant information should increase the saliency of the relevant

dimensions of positive instances which remain constant and thus facilitate discrimination.

Another point between these two studies or positions which may contribute to different findings is the type of stimulus display used when presenting experimental concepts. Bourne and Haygood (1961), Battig and Bourne (1961) used a distributed stimulus display with successive presentations where Haygood et al (1970), Chumbley et al (1971), used a more compact stimulus display with successive presentations. Compact stimulus displays differ from distributed displays in the way that the information pertinent to defining the concept is either presented on each trial or distributed over several trials where no one trial has enough information to define the concept.

Bourne and Haygood (1959, 1961) and Battig and Bourne (1961) presented positive and negative instances individually where Haygood et al (1970) and Chumbley et al (1971) presented a stimulus array of five stimuli (five letters) where each stimulus array was either a positive or negative instance. Given that Haygood et al (1970) and Chumbley et al (1971) used concepts which were defined by their relative position to the other stimuli in the five item array, presentation of such a stimulus array was necessary. Bourne and Haygood (1959, 1961) and Battig and Bourne (1961), used concepts which were defined by the combination of relevant and irrelevant information on critical and non critical attributes contained in one instance or noninstance of the concept class.

The differences between these two findings on intra dimensional variability may be situation specific to the type of relational concept and compact stimulus array used by Haygood et al (1970) and Chumbley et al (1971). Compact arrays allow subjects to scan more irrelevant information at one time reducing a memory load which is alternatively increased when using successive distributed stimulus arrays. Their results may have direct relevance to instructional programs which teach relationally based concepts presented in a compact stimulus display. Distar I (Language) uses a compact stimulus array when teaching relational concepts such as long and short. In their teaching sequence each stimulus set includes a number of examples for both concepts long and short. Instruction directs the student attention to each concept by showing its relation to the other opposite concept (i.e. this is long, this is short).

According to Haygood et al (1970) and Chumbley et al (1971) increasing irrelevant dimensions for long and short such as form, color, size and position with the relevant dimension of length remaining constant, should facilitate discrimination and identification of each concept. Becker et al (1975) suggest a cumulative programming strategy where irrelevant and relevant dimensions of a concept are introduced after previous discriminations are mastered, may not be necessary and may even be counter productive for relational concepts.

Concept identification research investigating the effects of stimulus complexity during original learning has not been

concerned with the influence of original learning conditions to generalization of concept usage. Generalization of concept usage is a major criteria for an operant definition of concept formation. The next section will review literature which investigated conditions of original learning and their effect on generalization of concept usage.

Conditions of Original Learning and Generalization

Another feature of concept learning not addressed by the concept identification research is the generalization from original learning conditions. The previously reviewed literature suggests intra and inter dimensional variability within relevant and irrelevant dimensions effects initial or original learning. However the effects of original learning were not assessed through generalization to novel instances.

The goal of concept teaching as stated by Becker, Englemann and Thomas (1971) is as follows:

> A concept has been taught when any or all members of the concept set are correctly identified (responded to in the same way) even though some were not in the teaching set and any or not all members of the concept set are responded to in a different way (p. 238).

This goal has an implicit measure or test to validate its effectiveness. The test is presentation of novel instances which were not observed during original learning. When a learner demonstrates generalization beyond original learning conditions the goal of concept teaching has been reached.

Stokes and Baer (1977) presented a review paper discussing the technology of generalization. They suggested that gen-

eralization is a "passive concept devoid of a technology" when compared to discrimination. Discrimination is a well understood process with an adequate technology which can be programmed and practised. Generalization, on the other hand, appears to occur without an adequate understanding of what "programmable" elements contributed to it. Stokes and Baer (1977) define generalization on pragmatic grounds,

> ... generalization will be considered to be the occurrence of relevant behavior under different, nontraining conditions (i.e. across subjects, settings, people, behaviors, and/or time) without the scheduling of the same events of those conditions as had been scheduled in the training conditions (p. 350).

An advancement of a technology of generalization would describe the initial training conditions which produced generalization as well as those which did not.

Stokes and Baer (1977) reviewed approximately two-hundred and fifty research studies categorizing each into one of seven categories. The one category which has relevance to the area of concept learning is "Training Sufficient Exemplars". As Stokes and Baer (1977) suggest, "in the training of sufficient exemplars, generalization to untrained stimulus conditions and to untrained responses is programmed by the training of sufficient exemplars (rather than all) of these stimulus conditions or responses" (p. 355). Their review suggested that the number of sufficient exemplars to produce generalization varied widely as a function of the behavior being trained. Different behaviors or classes of behaviors may represent instances which require different initial pro-

gram procedures to produce generalization. Stokes and Baer (1977) also suggest that in addition to "number of exemplars" a useful strategy may be to employ "diversity of exemplars". Diversity of exemplars although not clearly described, appears to mean inclusion of stimulus characteristics which change between different exemplars and represent stimulus characteristics of settings to which the behavior is to be generalized.

Becker et al (1975) introduced the concepts of interpolation and extrapolation when discussing the diversity of positive and negative instances for programming concept teaching sequences. Interpolation refers to a learner being able to identify new positive instances within the range of a concept when the previous training sequence samples instances representing a range of values defining the extremes of the concept. The learner learns the boundary values of a concept and any new concept instance within the previously taught range will be identified as a member of the concept.

Interpolation is concerned with identification of positive instances when teaching a limited number of examples from the range while extrapolation is concerned with identification of negative instances. Identification of negative instances results again from sampling. Here a learner learns the range of positive instances and that negative instances fall beyond the range. When learning that one instance is negative any new instance which falls further outside the range will be identified as a new negative instance.

When teaching characteristics of concepts which have a

range of values Becker et al (1975) suggest "sample the range with at least three instances but do not be exhaustive" p. 180. Two questions arise from Becker et al's (1975) suggestion namely, how does one arrive at a fixed number of instances to sample a range of values and secondly how different or diversified should the instances be. The issue of number of instances representing a range and the diversity of these instances within the range has not been clearly identified for individual concept classes. The principles of interpolation and extrapolation may be useful in programming sequences for teaching, however, as Stokes and Baer (1977) point out in their review the actual number and diversity of examples for adequate programming appears specific to individual concept classes.

Stokes and Baer (1977) suggest "the optimal combination of sufficient exemplars and sufficient diversity to yield the most valuable generalization is critically in need of analysis" p.357. Two research questions resulting from these two training conditions are stated by Stokes and Baer (1977):

1) is the best procedure to train many exemplars with little diversity at the outset and then expand the diversity to include dimensions of the desired generalization?
2) Or is it a more productive endeavor to train fewer exemplars that represent a greater diversity and persist in the training until generalization emerges?

The two dimensions of number and diversity of exemplars discussed by Stokes and Baer (1977) are similar to the dimen-

sions of intra and inter stimulus variability as described by Bourne (1961), Battig and Bourne (1961), Haygood et al (1970) and Chumbley et al (1971). Within concept identification tasks the interdimensional variability is the same as the diversity of exemplars, while intradimensional variability corresponds to number of exemplars.

As mentioned earlier the concept identification literature has concerned itself only with the role of these dimensions on initial learning and not to generalization. A logical step in the advancement of a technology for generalization within a stimulus class referred to as concepts would be to observe effects on generalization resulting from different stimulus dimensions employed in original learning.

Modigliani (1971) investigated the conservation of simple concepts as a function of the generality of the affirmation rule defining the concept. He taught sample concepts investigating the relation of age, and boundedness during acquisition to conservation of the concept under conditions where either the relevant or irrelevant dimensions were changed (transformations). Boundedness refers to the strength with which the defining attribute of a simple concept is bound to the irrelevant attribute. The term conservation in Modigliani's research is operationally the same as generalization of previous learned responses. Failure to conserve a response in the presence of new stimuli is the same as failure to generalize to new instances of a concept class. Stimulus materials were a flower in a flower pot where for

half the subjects the relevant dimension was shape of the leaves and for the others shape of the pot was relevant. The remaining dimensions of shape of blossom and orientation of stem were irrelevant.

The major interest of the study was the effects on conservation as a result of transformation on test items for either or both relevant and irrelevant dimensions. Transformations on the test items during original learning represented the test of generalization. Several transformations on original items for the generalization test were measured through a between group design. The various transformations or changes to original stimuli were as follows: 1) substitution of irrelevant dimensions 2) addition of new irrelevant dimensions, 3) deletion of irrelevant dimensions 4) combination of deleting all irrelevant dimensions and adding new ones along relevant dimensions. The purpose of these transformations was to assess their influence on generalization.

The results of specific interest to the present review are the addition and deletion of relevant and irrelevant information. There was no significant difference between adding one versus many irrelevant dimensions, however deletion of all irrelevant dimensions was the most effective in producing non-generalizing responses. Addition of relevant information (different but similar) resulted in a decrease in generalization which was similar to deletion of all irrelevant information and significantly greater than addition of irrelevant information.

Modigliani's (1971) results indicate that in generalization of the affirmation rule, irrelevant information has psychological significance. Conservation less than 100% of the affirmation rule was affected by the following transformations (increasing order of decrement); a) substitution of one original irrelevant vale; b) change in the number of irrelevant values, either by addition or deletion of old ones; and c) radical change in context (addition of relevant dimensions).

Bourne (1969) and Hunt (1962) state that once a rule is learned it will be generalized to all stimuli demonstrating the defining value. Modigliani (1971) has shown that in the presence of the cue defining a concept, generalization is governed by the nature of both relevant and irrelevant information. He concluded that "the presence of a defining attribute is a necessary but not sufficient condition for the identification of a stimulus as an instance of the concept" p. 239.

Modigliani and Rizza (1971) further studied the deletion of irrelevant attributes on the generalization of simple concepts. They suggested that subjects chose one of two alternative strategies when acquiring rules for concept generalization. One strategy shows demonstration of a general rule (G) which suggests that the defining value is the necessary and sufficient condition for new stimulus to be identified as a concept instance during a test of generalization. The second strategy demonstrates a within-set rule (WS), a rule

which restricts the conceptual category and resultant generalization to that subset of training stimuli containing the defining value e.g. (total context of relevant and irrelevant dimensions). For the WS rule irrelevant information plays a significant role limiting the amount of generalization whereas the G rule does not. Their results showed that subjects may adopt one of two rules during acquisition, however, both result in different levels of conservation or generalization.

As Modigliani and Rizza (1971) point out, their test of novel items is not the standard test of generalization. Their test involved the deletion of irrelevant dimensions whereas most tests would involve items which vary on the same number of intra or inter dimensional variability (i.e. Bourne, 1961).

Bourne (1961), Bourne (1965), Bourne, Ekstranck and Dominowski (1971), suggested:

> One of the trademarks of concept learning is that once the concept has been mastered for a large set of stimulus instances, novel stimuli elicit the correct response, positive or negative, without hesitation or error by the subject (p. 195).

Given the work of Bourne (1961) and Battig and Bourne (1961), one would expect a greater number of training trials to criterion as the within stimulus variability increases. However, as Modigliani (1971), and Modigliani and Rizza (1971) point out, there is no reason to assume that subjects will generalize to instances which do not contain the same number of irrelevant dimensions during training. Stokes and Baer (1977) suggest that generalization may be a function of the

number of stimulus characteristics during training. Training with a large number of exemplars and high level of exemplar diversity may result in not only more training trials but also generalization limited to the training conditions.

The training context or stimulus arrangement may well limit the extent of generalization. The studies reviewed by Stokes and Baer (1977), all employed a cumulative program design where in the absence of generalization, new conditions were added until generalization emerged. Cumulative programming and task sequencing is advocated by Becker et al (1975), to counteract the teaching of misrules.

A cumulative approach to introducing new information in a concept learning situation should result in generalization to new concept instances regardless of the transformation or changes made in the type or number of relevant and irrelevant dimensions. This would be due to persons learning situations where all combinations of number and type of irrelevant and relevant dimensions were experienced and mastered by the learner. There does not appear to be any experimental evidence supporting the notion that cumulative programming will counteract the learning of misrules or lack of generalization.

Summary

The role of original learning conditions on generalization is not clear. Stokes and Baer (1977), provide a review of literature which shows that cumulative programming in the absence of generalization results in an increased ease when

learning behaviors in a new context and that generalization will eventually emerge. Becker et al (1975), provide a logical, although not empirically demonstrated, explanation for cumulative programming to facilitate generalization.

It is still not clear how original learning conditions affect a learner's use of interpolation and extrapolation when confronted with new stimulus material. A second issue concerning original learning conditions, as Stokes and Baer (1977), point out, is the question of the most optimal training conditions to increase generalization.

The previous review of literature dealing with conditions of stimulus complexity during initial acquisition of concept identification discussed the effects of inter and intra dimensional variations for relevant and irrelevant information and how they affect acquisition. The comment by Haygood et al (1971) and Stokes and Baer (1977), that initial training conditions and their effect on generalization may be specific to different concept classes is very appropriate. The rule defining the concept class and mode of presentation dictated by the rule may contribute to initial training condition differences on generalization.

For example, Bourne and Haygood (1961), and Battig and Bourne (1961), found generally, that increases in stimulus complexity resulted in increased trials to solution, however Haygood et al (1971) and Chumbley et al (1971), found the opposite result. The difference between these two findings may be due as Haygood et al (1971) suggest, to the nature of

the concept (relational in their study and disjunctive in that of Bourne and Haygood, 1961) and the manner in which stimulus complexity was increased (only irrelevant dimensions for Haygood et al (1971), and both relevant and irrelevant for Battig and Bourne (1961).

As mentioned earlier, generalization from original learning was not assessed and the only clue to a generalization effect may come from the research of Modigliani (1971). In absence of cumulative programming, Modigliani (1971), found that when adding new cues on old dimensions or deleting irrelevant cues from the original learning context resulted in deterioration of generalization. Presently the knowledge on experimental results on original learning conditions and generalization is difficult to compare directly as different concept classes were used with different independent variables being manipulated.

Given the conclusion by Stokes and Baer (1977) that different behaviors may require different antecedent programming for generalization it would appear appropriate to further investigate one concept class more thoroughly, isolating the program characteristics which influence generalization. The present research will investigate the role of stimulus complexity for increasing irrelevant information in relational polar opposite concrete concepts presented in a compact stimulus array and the effects of different levels of stimulus complexity on generalization. The concepts used, stimulus presentation format and instructional sequence used

closely resemble those used in educational programs to teach concepts to preschool mentally retarded children. Using this format, the experimental results will have direct application to program considerations in educational planning.

Applied Research on Concept Teaching with Preschool Mentally Retarded Children

Bellamy and Bellamy (1974), present a study investigating concept learning with naturally occurring concrete concepts and a subject population consisting of handicapped children. Their population included preschool mentally retarded children enrolled in a preschool program for the retarded. The concepts taught were called descriptive concepts and represented polar opposite descriptive adjectives (big-little, hot-cold, long-short, straight-curved). Instruction took place in a small group setting, however, it was not stated how many children were in the group.

The testing and teaching procedures consisted of a baseline condition followed by two concepts being taught to criterion and a probe set for generalization. Concepts were pre- and post-tested and taught in pairs. Each concept showed generalization to novel instances following criterion during training. This study did not investigate analytically any dimensions of concept teaching, however, it did suggest that; one, their teaching procedures were sufficient to develop acquisition of concepts presented in pairs using a simultaneous presentation format and secondly, acquisition resulted in generalization to novel instances of the concept class

taught. The teaching procedure realized the goal of concept teaching as stated by Becker, Englemann and Thomas (1971).

Teaching and testing sessions were arranged into trials. During a trial each student made six responses to three object pairs representing the concept pair being taught. Within a trial each object pair was presented twice and concepts representing each object pair were taught simultaneously. On one presentation one concept from the pair was requested, while on the second, the other concept was requested.

The instructor requested a nonverbal touching response from the children following the instruction "touch (name of concept)". The stimulus materials were either real objects or two dimensional pictures. The stimulus materials for each teaching and probe set had the same inter-dimensional stimulus variability. In all instances of object pair presentation only the relevant defining dimension was different, all irrelevant dimensions were constant (i.e. big, balloon, red, 4" diameter, little, balloon, red 2" diameter). The teaching sets for each concept pair consisted of three stimulus pairs. Each stimulus pair was different on all dimensions representing a new example of the concept (i.e. concept big-little; balls, circles, hershey chocolate bars).

The test for generalization consisted of novel stimulus items presented in pairs with the same inter-stimulus variability observed during teaching. As the relevant dimension defining each concept was not varied with other possible inter stimulus variabilities on irrelevant dimensions, it is possible

the student would not have generalized to novel instances with different levels of inter-stimulus variability.

The concept used in this study had only one relevant dimension which was required for definition. The dimension was either size, temperature or length. Each concept member of a pair was defined by its relation to the other member of the pair along one of these dimensions. It was this relation as a defined rule which forms the basis for generalization to novel instances.

The role of inter-stimulus variability or diversity of exemplars was not systematically varied in this study leaving the possible conclusion that misrules or lack of generalization could result. Any deviation in the generalization test for within item variability from addition of irrelevant dimensions may have resulted in a lack of generalization. The research of Modigliani (1971) and Modigliani and Rizza (1971) although working with different concepts and presentation format would predict this.

The role of number of exemplars as a factor in acquisition and generalization of concept usage was not evaluated in the present study as it was held constant for each concept teaching set. The study does suggest that three exemplars with inter-dimensional variability for irrelevant dimensions held constant, results in generalization to novel instances. However, given the degree of inter-stimulus variability, possibly one or two exemplars could have been enough to produce the same level of generalization within the same or few-

er number of teaching sessions.

There are a number of unanswered questions about the stimulus variability during training and resultant generalization from this study. In addition to the validation of stated objectives the study demonstrated that preschool retarded children can acquire and generalize concept usage.

Relational Concepts

There is a lack of research investigating the stimulus characteristics involved during acquisition of relational concepts and resultant generalization. Relational concepts are defined by stimulus cues which are in relation to each other. The critical or defining attribute suggests a relation between stimulus cues along a specified dimension. Dimensions for relational concepts are represented in quantity (i.e. more-less) size (i.e. big-little), time (i.e. before-after) or spatially (i.e. in-out).

One class of relational concepts are polar opposites. Polar opposite concepts vary as extremes along a cue dimension. Rules defining the concept attribute are usually unidimensional where only one attribute is required to describe instances and noninstances. Generalization to novel instances demonstrates use of the affirmation rule.

There are a number of instructional programs designed to teach children polar opposite concepts (Distar, 1971; Fredrick et al, 1976; Anderson et al, 1977). With the exception of Distar, these programs do not describe the stimu-

lus characteristics to employ during instruction. All these programs suggest a simultaneous presentation where either one or multiple instances of both polar opposite concepts appear together. One member of the polar opposite pair acts as a positive instance of itself and a noninstance of the opposite concept. This presentation format relies heavily on the relation between instances and noninstances.

It is not clear what role irrelevant information plays in altering a subject's attention to relevant dimensions for polar opposite concepts during acquisition and generalization of the rule. Relational rules and more specifically polar opposites presented in a simultaneous format have not previously been investigated under the task dimensions previously discussed.

The research of Bourne (1961) and Battig and Bourne (1961) would suggest a decrease in performance during acquisition as the number of irrelevant dimensions increase. Haygood et al (1970) would suggest that as intra-dimensional variability increases, where concept complexity does not increase, positive performance during acquisition should increase. In a situation where the rule of difficulty does not increase and irrelevant information is increased it is possible to predict an increase in performance over trials due to saliency of relevant dimensions (Archer, 1962).

Haygood et al (1970), pointed out that rules which place emphasis on the negative instances for confirmation of positive instances may show a decrease in performance as irrele-

vant information is increased. Relational concepts presented simultaneously will rely on the relational rule where instances are compared to noninstances. In such a situation, increases in irrelevant information may result in decreased performance over time.

Summary

Generalization of relational concept rule usage has not previously been investigated as a function of stimulus variability during original learning. Bellamy and Bellamy (1974) demonstrated that preschool retarded children acquire and generalize relational concept rule usage with polar opposite descriptive adjectives. The generalization test demonstrated rule usage to novel stimulus items pairs with similar inter-dimensional variability. However, as Bellamy and Bellamy (1974) used a test of generalization which only included test pairs items with inter-dimensional variation observed during initial teaching and a non-cumulative teaching program, it is highly probable that the children in their study would not have demonstrated the same level of generalization to novel instances displaying inter-dimensional variation not incorporated in the training setting. The confusion over stimulus complexity and the role of irrelevant dimensions for relational concrete concepts during initial learning and generalization is in need of further clarification as it applies to educational programming for the mentally retarded.

CHAPTER III

HYPOTHESIS

General Rationale

The previously reviewed literature centered on the role of stimulus complexity during initial learning situations and generalization of concept usage. Due to a lack of experimentation, there does not appear to be a clear understanding of how stimulus complexity in concept teaching situations influences generalization. In addition there is some confusion over the influence of stimulus complexity during initial learning for different concept classes, and presentation modes employed.

As stimulus complexity during initial learning for optimal generalization may be specific to individual concept classes, the present research will attempt to investigate two parameters of stimulus complexity with a specific class of concepts. Concept teaching programs have been designed and packaged as curriculums for teaching the mentally retarded children. There is an undesirable lack of curriculum validation and demonstration of effectiveness with the populations for which these programs are designed. Several of these programs do not describe the nature of the stimulus materials to use during teaching, in addition to not defining adequate procedures for an instructor to use when monitoring the effectiveness of the instruc-

tional program. Although general guidelines are usually outlined there is a lack of specification concerning the nature and type of instructional materials.

The present research has attempted to focus on two issues arising from the previously reviewed experimental and applied research in concept learning. The first is concerned with investigating the role of stimulus complexity during initial learning and generalization in an attempt to further a technology of generalization. This was done through isolating different programmable dimensions of stimulus complexity and observing the amount and degree of generalization which occurs.

The second concerns instructional procedures when teaching concepts to preschool mentally retarded children. Due to a lack of published information for concept teaching programs for this population, the present research may provide valuable descriptive data concerning acquisition and generalization for the class of concepts taught.

The purpose of this research then, is an investigation of stimulus complexity as one task dimension in the initial learning and generalization of concepts with pre-school mentally retarded children. More specifically the study will assess the effects of two stimulus parameters during initial learning on generalization. These two stimulus parameters are inter and intra stimulus cue variability which parallel the two dimensions of "number" and "diversity" of exemplars discussed by Stokes and Baer (1977).

The concept training task will employ relational concepts

(polar opposites) taught in a simultaneous format within a receptive paradigm. Polar opposition concepts represent a class of relational concepts with one relevant cue dimension and a large number of possible irrelevant dimensions.

The two instructional parameters, "number" and "diversity" of exemplars were tested for their independent and interactive effects within a two by two factorial design. Four experimental conditions resulted which tested two levels of number of exemplars (two versus four) and two levels of exemplar diversity (one irrelevant dimension versus three irrelevant dimensions). The dependent measure during acquisition for each concept trained was number of trials to criterion.

Generalization was tested following acquisition of concepts within an experimental condition. Generalization was assessed at three different generalization tasks each containing a different level of stimulus diversity. The dependent measure during generalization tests was number and percentage of generalization responses.

<u>Definitions</u>

This section is designed to further describe each stimulus dimension under investigation in addition to definitions within the study.

<u>Number of Exemplars (NE)</u>

Number of exemplars refers to how many exemplars representing a concept are included in a teaching set for an experimental condition. For example, if you were teaching the concept "ball" a teacher may use two or more different balls. The number of

stimuli sampled from the possible range of stimuli representing a concept is referred to as number of exemplars.

The present research investigated two levels of number of exemplars. Number of exemplars level one included teaching sets where two stimulus items represented the concept taught. Number of exemplars level two, included teaching sets where four stimulus items represented the concept taught. The difference between levels one and two is only in the number of stimulus items which were used during instruction.

Interdimensional Stimulus Variation (IER)

Interdimensional stimulus variation corresponds to the term diversity of exemplars discussed by Stokes and Baer (1977). These two terms can be used interchangably. Interdimensional stimulus variation (IER) and diversity of exemplars refers to the number of relevant and irrelevant dimensions which are different or similar between stimulus items.

The present research study presented polar opposite concepts simultaneously in a receptive paradigm. Therefore, during instruction a child was presented with both members of a polar opposite concept and was requested to choose one member. In the present study then, interdimensional stimulus variation (IER) refers to the number of relevant and irrelevant dimensions which were similar or different between stimulus items representing polar opposite concepts.

There were two levels of interdimensional stimulus variation (IER). Level one IER represented concept instances which were different on the relevant dimension defining the concepts

while all irrelevant dimensions were the same. For example if the concept pair was "long-short", a child would see "long" as a green pencil eight inches in length and "short" as a green pencil four inches in length. The relevant cue length was different, however the irrelevant cues color and form were similar.

Level two IER represented concept instances which were different on the relevant dimension defining the concept and all other irrelevant dimensions (color, form). An example of level two IER for the concept pair "long and short" would be "long" presented as a green pencil eight inches in length with "short" as a red wagon four inches in length. Here the relevant cue for individual concept identification is different while all other irrelevant dimensions are also different.

Generalization

Generalization for the present study was defined as correct responses to novel exemplars of the concept class which were not present during instruction. Generalization was assessed at three different generalization tasks.

Generalization task one (G.T.1) consisted of stimulus items of a concept pair presented simultaneously where the relevant dimension defining the concepts was different and all 3 remaining irrelevant dimensions were the same. For example a generalization test pair for "straight-curved" would contain a purple rope eight inches in length ("straight") and a purple rope eight inches in length curved into a half circle shape ("curved"). The relevant dimension defining the

concepts "straight or curved", was different, while color and form as irrelevant dimensions, were the same.

Generalization task two (G.T.2) contained items where the relevant dimension defining the concept members was different while one irrelevant dimension (color), was different and the remaining irrelevant dimensions were the same (form). An example for "straight-curved" at G.T.2 would contain a stimulus pair with a straight purple rope eight inches in length and a curved black rope eight inches in length. The relevant dimension curvature and irrelevant dimension color are different while the remaining irrelevant dimension form is the same.

Generalization task three (G.T.3) contained items where the relevant dimension defining the concept members was different and all remaining irrelevant dimensions (color and form) were different. An example for "straight-curved" at G.T.3 would contain a stimulus pair with a straight brown coat rack (straight) eight inches tall, and a yellow half moon five inches from tip to tip (curved). Here, curvature as a relevant dimension is different, while irrelevant dimensions color and form are also different.

Hypotheses and Rationale

General Hypothesis One

As the interdimensional stimulus variation increases between concept pairs for irrelevant dimensions there will be an increase in the number of trials to criterion during initial learning.

Specific Hypothesis

The mean number of trials to criterion for IER level one will be less than IER level two at each level of number of exemplars.

General Hypothesis Two

As the number of exemplars (NE) for each concept pair increases there will be an increase number of trials to criterion during initial learning.

Specific Hypothesis

The mean number of trials to criterion for NE level one will be less than NE level two at each level of interdimensional stimulus variation.

Rationale for Hypotheses One and Two

The previous research of Bourne (1961) and Battig and Bourne (1961) has shown that as stimulus complexity increases due to inter versus intra stimulus variation there is also an increase in trials to problem solution. Haygood et al., (1971) and Chumbley et al., (1971) have shown the opposite result when increasing intra stimulus variation with conjunctive concepts with stimulus relations as a defining rule. These two studies suggested, however, that this result may not be the same for concepts which place emphasis on negative instances for identifications. Relational concept classes such as polar opposites presented in a simultaneous format do place emphasis on negative instances for identification of positive instances. For this reason it is hypothesized that regardless of the type of stimulus variation the results of

the present research will be consistent with the findings of Bourne (1961) and Battig and Bourne (1961).

General Hypothesis Three

As the number of exemplars (NE) from a concept class increased during training (two versus four) there will be an increase in the total generalization responses following acquisition.

Specific Hypothesis

The average number of total generalization responses will be greater for NE level two over NE level one at each level of interdimensional stimulus variation.

General Hypothesis Four

As the number of exemplars (NE) from a concept class increased during training (two versus four) there will be a greater increase in total number of generalization responses at generalization task one, two and three.

Specific Hypothesis

The average number of generalization responses at generalization task one, two and three will increase as the number of exemplars increases from two to four at each level of interdimensional stimulus variation.

Rationale for Hypotheses Three and Four

Stokes and Baer (1977) reviewed a number of research articles which suggested that generalization increases as a function of the number of exemplars during initial training. These studies all used a cumulative programming approach where new training examples were gradually introduced until

generalization emerged. The present research is not using a cumulative programming approach but rather different numbers of exemplars during initial teaching. It is hypothesized that this difference in initial training conditions will have a similar result in generalization as demonstrated by cumulative programming.

General Hypothesis Five

As the interdimensional stimulus (IER) variation increases between concept pairs for irrelevant dimensions there will be an increase in the total generalization responses following acquisition.

Specific Hypothesis

The average number of total generalization responses will be greater for IER level two over IER level one at each level of number of exemplars.

General Hypothesis Six

As the interdimensional stimulus variation increases between concept pairs for irrelevant dimensions there will be an increase in the average total number of generalization responses at each generalization task - one, two and three.

Specific Hypothesis

The average number of generalization responses at each generalization task (tasks one, two, three) will increase as the interdimensional stimulus variation increases from level one (one irrelevant dimension) to level two (three irrelevant dimensions) at each level of number of exemplars.

Rationale for Hypotheses Five and Six

Modigliani (1971), Modigliani and Rizza (1971) found that changes in irrelevant features on generalization tests resulted in a decrement in performance. Their initial training conditions for the concepts did not employ a cumulative programming sequence. The lack of variation in irrelevant dimensions and variation in combination with relevant dimensions appears to reduce the possibility of generalization when irrelevant dimensions are changed in novel tasks.

Modigliani (1971) concluded that subjects were responding to novel instances not only with attention to relevant dimensions but also to irrelevant dimensions. Therefore learning a concept as a total context with both relevant and irrelevant dimensions resulted in poor generalization when irrelevant dimensions changed while relevant dimensions remained constant.

Given the non-cumulative approach to training during initial learning in the present experiment it is hypothesized that generalization responses will be specific to the same level of IER variation observed during initial learning.

The next chapter describes the individual children who acted as subjects and general procedures employed in the present study.

CHAPTER IV

METHOD

Subjects

Two moderately retarded boys, one moderately retarded girl and one severely retarded boy served as subjects. One moderately retarded boy was enrolled in the Early Education Project (Kysela et al., 1977) located in Mayfield Elementary School. This child had been in the project for a period of three years. During this time he had learned a number of self-help, language and general cognitive behaviors. The remaining three children were enrolled in the early childhood classes at Winnifred Stewart School. Winnifred Stewart School is a private school for the education of educable and trainable retarded persons. These three children had been enrolled in the early childhood class for one year where they had learned a number of self-help, language and general cognitive behaviors. One child, a severely retarded boy, was receiving individual speech therapy sessions four times a week.

These children were selected for the experiment as they demonstrated the following characteristics: a) chronological age which did not exceed six years; b) developmental delay of approximately one year or more as assessed by a standardized intelligence test; c) they were able to pay attention to the experimenter for a prescribed period of time; d) they could

perform a touching response; e) they failed to exceed the chance level in identifying members of a concept class during a baseline test. Table one summarizes the psychometric characteristics of the children and Table three shows the results from the baseline assessment.

Apparatus

The stimulus materials consisted of two dimensional pictures of objects depicting relational concepts. These pictures were of objects common to most natural environments rather than abstract or experimentally defined concepts. The relational concept pairs employed were; "long-short", "in-out", "straight-curved", and "empty-full". Stimulus materials representing each member of a concept pair were reproduced on a rectangular card eight inches by six inches. Appendix B describes examples of stimulus materials in each experimental condition.

Experimental Design

The present research employed a within-subject-design (Sidman, 1960). This design, in a multiple baseline fashion, demonstrates experimental control of independent variables through the use of baseline logic and replication of experimental conditions across experimental subjects (Kazdin, 1976).

There were four experimental conditions in the experiment. Each experimental condition contained two instructional parameters for task presentation which have been postulated to influence generalization of learned responses. These two para-

Table I

Diagnostic and Psychological Characteristics of the Children

Children	Diagnosis	Chronological Age Years	Chronological Age Months	Mental Age Years	Mental Age Months	I.Q. at Last Testing
Child One	Down's Syndrome	5	6	4	2	Standford Binet = 66
Child Two	Down's Syndrome	4	1	3	0	Standford Binet = 62
Child Three	Mental Retar- dation. Etiology unknown	4	1	3	1	Standford Binet = 64
Child Four	Down's Syndrome	5	3	2	0	Unknown* < 50

* I.Q. estimated from Expressive Language Survey, referred to in school file as untestable.

meters are, number of exemplars and diversity of exemplars. These two parameters were separated into two levels resulting in four experimental conditions. The four experimental conditions described below were designed to investigate the main and interactive effects for number and diversity of exemplars during original learning and generalization.

Experimental Conditions

Table two describes how these two parameters and level of each are combined into the four experimental conditions.

Number of Exemplars (NE)

This parameter had two levels for presenting concept pairs in a teaching set. Level one has two different exemplars of the concept pair, while level two has four. For example, number of exemplars level one, for the teaching set with concept pair "long-short" contained a long and short pencil (exemplar one) and a long and short wagon (exemplar two). Level two number of exemplars for the same concept pair contained a long and short pencil (exemplar one), long and short wagon (exemplar two), long and short train (exemplar three) and a long and short fish (exemplar four).

Diversity of Exemplars (IER)

This parameter refers to the degree of interstimulus variability between concept members for each exemplar. There are two levels of interstimulus variability (IER). Level one has one irrelevant dimension (position) varied between members of a concept pair, while level two has three irrelevant dimen-

Table II

Diversity of Exemplars (IER)

Number of Exemplars (NE)		
	Condition ONE a) Teaching set contains two pairs of exemplars, level 1 NE. b) Within pair difference on one irrelevant dimenstion (position), level 1 IER. c) Within pair similarity on remaining irrelvant dimensions (form/relative size, color).	Condition Three a) Teaching set containts two pairs of exemplars, level 1 N.E. b) Within pair difference on three irrelevant dimensions (position, form/relative size, color), level 2 IER.
	Condition Two a) Teaching set contains four pairs of exemplars, level 2 NE. b) Within pair difference on one irrelevant dimension (position), level 1 IER. c) Within pair similarity on remaining irrelevant dimensions (form/relative size, color).	Condition Four a) Teaching set contains four pairs of exemplars, level 2 NE. b) Within pair difference on three irrelevant dimensions (position, form/relative size, color), level 2 IER.

sions (position, form and color), varied between members of a concept pair.

For example, at interstimulus variability level one for the concept pair "long-short" the stimulus materials contain a green pencil eight inches long ("long") and a green pencil four inches long ("short"). At interstimulus variability level two the stimulus material would contain a green pencil eight inches long ("long") and a red wagon four inches long ("short").

During instructional teaching sets the irrelevant dimensions are varied between both members of the concept pair. Using the previous example of green pencil eight inches long ("long") and a red wagon four inches long ("short"); a teaching set would contain a complete counterbalancing for all irrelevant dimensions. There would be exemplars for "short" in both the right and left position which were red and green, pencils and wagons four inches long, in addition to exemplars for "long" in both the right and left positions which were red and green, pencils and wagons eight inches long.

Procedure

Each child was introduced to the following conditions; pre-baseline and adaptation period, baseline condition, experimental and probe condition and generalization test condition.

Pre Baseline and Adaptation Period

The purpose of this condition was to ensure that each

child has the necessary prerequisite behaviors for the experiment in addition to reducing possible reactive effects to the experimenter and experimental setting. The defined prerequisite skills were subject attention to the experimenter upon verbal instruction and demonstration of a nonverbal "touching" response to previously known items following the experimenter's verbal instruction. Each of these prerequisite skills are operationally defined and stated in behavioral terms according to the format adopted by Sulzer-Azaroff and Mayer (1977).

Attention: Goal: Subject will look at the experimenter upon request.

Situation or Conditions: One to one teaching situation with experimenter and child seated at age appropriate table and chairs.

Criterion level: Respond correctly 80%.

Behavioral Dimensions: Frequency: At least eight out of ten times during a ten minute period.

Intensity: Not applicable.

Topography: Child attending response is defined as orientating their eyes and head toward the experimenter to produce eye contact.

Duration: A correct response will be eye contact for a minimum of two seconds.

Latency: The child will initiate his attending response within a two second period following experimenter command.

Behavioral Objective: Child will look at the experimenter and engage in eye contact for a minimum of two seconds within two seconds following the instruction, "(child's name) _____ look at me", eight out of ten times.

"Touching Response: Goal: Child will touch familiar objects and two dimensionsal pictures upon experimenter request.

Situation or Conditions: One to one teaching situation with the child seated at an appropriately sized table and chair with the experimenter at the opposite side of the table; variety of familiar objects and pictures.

Criterion Level: Respond correctly 80% to both objects and pictures.

Behavioral Dimensions:

Frequency: At least eight out of ten times during a ten minute period for objects and pictures.

Topography: Child correct responses are defined as extension of one arm and hand from a resting position on the table or lap, and physical contact with the object or picture with any portion of his hand.

Duration: A correct response will have to be completed within five seconds following subject initiation.

Latency: The child will initiate the response within two seconds following the experimenter command.

Behavioral Objective: Given a set of five objects (three dimensional) and five two dimension pictures each presented twice in a random alternating series, the child will initiate touching the object or picture within two seconds of the

experimenter's request "(child's name) _____, touch (object or picture name) _____," and complete the touching response within five seconds of initiation. The minimum level of performance for each set of objects and pictures where each is presented twice is 80% correct.

Pre-school children and especially mentally retarded children can display a negative reaction to testing or initial teaching situations which are novel or unnatural for them. As a result a child's performance may not be a true and reliable index of his ability (Baine, 1977). In an attempt to reduce "reactive effects" on performance due to novel instructor, setting, mode of stimulus presentation and format/or instruction, each child was introduced to one fifteen minute adaptation session.

During the adaptation session the experimenter took each child to the experimental setting. Stimulus items of similar dimensions to those in the experiment were presented in a two choice discrimination format. Each child was requested to touch the appropriate stimuli upon request. The list of stimulus materials used during the adaptation session contained two novel items and several items with which individual children had prior experience. The selection of previously known items was obtained through consultation with each child's teacher. The two novel stimulus pictures were included to introduce the child to the correction procedure. These items were not used in the experiment.

Throughout the adaptation session each child was socially reinforced for sitting and attending to the instructor on a variable ratio schedule (VR3).

Baseline Condition

During a baseline condition each child's knowledge of the four concept pairs or eight concepts was assessed. Each concept was assessed 18 times by nine different pairs of exemplars randomly located in both the right and left position. There were three different baseline tasks under which each concept member was assessed.

Baseline task one contained three different exemplars where the interstimulus variability between concept members varied on only one irrelevant dimension (position). Task two contained three different exemplars where the interstimulus variability between concept members varied on two irrelevant dimensions (color and position), while task three contained three different exemplars with interstimulus variability between concept members on three irrelevant dimensions (position, color and form). Items for the baseline test were selected randomly from the pool of generalization items which appears in Appendix D. There was a total of 144 trials required to assess all the concept pair members six times on each baseline task. The baseline test was spread over two individually administered twenty minute sessions.

Table three shows the results of the baseline test for each child. Each child satisfied the criterion for inclusion in the experiment, of less than, or equal to fifty percent

Table III

Baseline Assessment

Children	Concept Pair Long	Concept Pair Short	Concept Pair In	Concept Pair Out	Concept Pair Empty	Concept Pair Full	Concept Pair Straight	Concept Pair Curved
	Level One		Level One		Level One		Level One	
Child one	50	16	16	33	33	16	16	33
Child two	50	50	50	16	50	50	50	50
Child three	0	50	33	33	33	33	33	33
Child four	33	16	16	33	33	16	16	16
	Level Two		Level Two		Level Two		Level Two	
Child one	0	0	33	33	33	16	33	33
Child two	0	16	50	16	50	50	33	0
Child three	50	50	33	16	33	0	33	33
Child four	16	0	16	16	16	16	16	0
	Level Three		Level Three		Level Thre		Level Three	
Child one	33	16	16	20	33	16	16	33
Child two	0	16	33	16	0	33	33	0
Child three	16	16	16	30	33	0	16	33
Child four	16	16	0	16	33	16	0	0
	Total		Total		Total		Total	
Child one	28	11	22	16	33	16	22	33
Child two	17	27	44	33	33	33	38	16
Child three	22	39	27	33	33	16	27	33
Child four	22	11	11	22	27	16	11	6

correct responses on each baseline task and total presentations. It is interesting to note how children 3 and 4 responded almost consistently with a level of fifty percent correct for both totals and separate baseline tasks. These children both demonstrated a position response set to the right position. Their baseline results are predictable as the items were counterbalanced for position. Children 1 and 2 demonstrate a more random alternating strategy. There may be some relationship between position responding for subjects 3 and 4 during baseline and acquisition phases as they required a greater number of sessions to criterion than children 1 and 2. Initially during acquisition a large number of trials may have been spent correcting this response set.

Procedure During Baseline Condition

Stimulus items representing concept pairs were presented in front of the children in a two choice discrimination format. Testing on each trial consisted of the experimenter initially gaining the child's attention and then requesting him to touch or point to a stimulus card representing the concept named (i.e. "Touch ___"; or "Point to _____.") During each trial only one member of the concept pair was tested. This concept pair was presented at a later time in the random sequence to test the opposite member.

The experimenter did not consequate a child's response to baseline items, however social praise for sitting or attending to the instructor was given on a VR3 schedule. As a child's pointing or touch response was not consequated

a series of competency items were inserted within the baseline series approximately every seventh item. These competency items taken from the list of known items used during the adaptation session were consequated by the experimenter. This procedure ensured an intermittent schedule of feedback to the child for their touching or pointing response. This was done to reinforce on-task behavior in addition to the response characteristic demanded by the task. These competency items did not contain examples of any of the concepts used in the experiment. Such a procedure, with a similar rationale, has been used previously by Paynon and Hall (1977).

Instructional Procedure for Experimental Conditions

A pilot study completed with two children prior to the main experiment revealed a flaw in the original instructional procedure. The original procedure consisted of presenting both members of a concept pair simultaneously and first requesting one member and secondly requesting the opposite member without changing their position or interchanging new stimulus cards.

This procedure created a response set for both children, where they automatically pointed to the second card before any instruction and immediately following feedback on their first response. The presence of their response set was demonstrated in an A-B-A design. The pilot study and analysis of the response set are presented in Appendix A.

The instructional procedure was subsequently changed for the main experiment. The new procedure consisted of presenting both members of a concept pair simultaneously and only one concept was requested during the trial. These stimulus items appeared again later in the series of trials and the opposite concept member was requested. The instructional sequence consisted of initially securing the child's attention and then presenting an instructional trial. If a child did not attend upon request by either orientating himself toward the experimenter or stimulus cards, a step-down remedial procedure was initiated (Kysela et al., 1977). This procedure consists of introducing additional verbal and gestural prompts and hands on guidance in a cumulative manner until the child attends. Correct responses to the attention signal were followed by social praise on an intermittent schedule.

Correct responses by each child on the instructional trial were followed immediately by praise feedback, descriptive praise and for one subject, consumable reinforcers. Incorrect responses to initial requests were followed immediately with negative feedback and a retrial procedure. Negative feedback consisted of saying, "No" or "Wrong", while the retrial procedure consisted of a modelling of the correct response by the experimenter ("This one is long", experimenter touches it) and another request to touch the named concept ("You touch ____"). Incorrect responses following the retrial procedure were ignored.

Figure 1 shows a flow chart of the instructional procedure.

Teaching sessions completed on a one-to-one ratio were approximately twenty-five minutes in length. Individual sessions for two children were completed in the counsellor's office at Mayfield Elementary School (10' by 20'). The other two children attending Winnifred Stewart School were given sessions in a storage office (10' by 30'). These teaching environments were relatively free of competing visual and auditory stimulation.

Two experimental conditions were taught during each session. They were not taught simultaneously, but rather in two separate blocks of 32 trials. The two conditions were alternated each day to counterbalance the order of teaching blocks. Each condition contained 32 trials, 16 trials for each member of the concept pair. The arrangement of the stimulus items in instructional series for concept pairs with each experimental condition is described in Appendix C.

The criterion for acquisition of a concept pair was 80 percent or greater correct responses of the total during one complete teaching session. One member of a concept pair may reach 80 percent correct responses before the other, however, the criterion for acquisition is 80 percent or greater correct responses for both members of the concept pair within one session.

If a child did not reach this criterion for both members of a concept pair within 250 trials instruction was terminated.

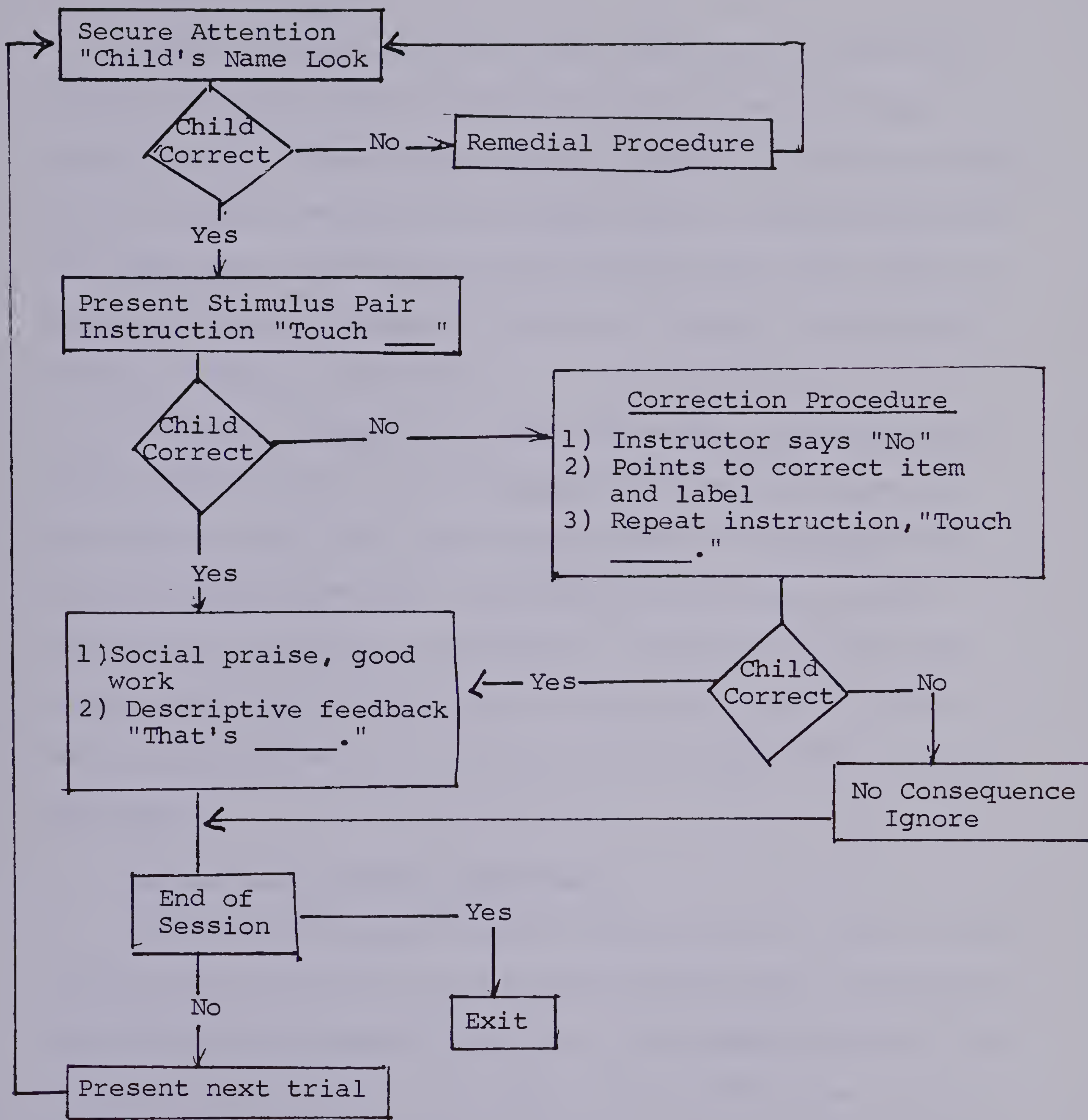


Figure 1. Flowchart of Instructional Procedure

Probe Condition

During teaching sets for both conditions a series of eight probe items were interspersed approximately every fourth trial. These probe trials tested a child's knowledge of the concept members to be taught next. Probes consisted of items with interstimulus variability level two, that is, three irrelevant dimensions (position, color, form) varied between members in the pair.

A child's response to probe items was not consequated by the experimenter. This procedure of probing items to be taught in the next experimental condition indicated if concept members had been learned following the original baseline and prior to instruction. A probe item was considered learned if there was a total of 80 percent correct across all probes or 80 percent correct over three teaching sessions.

Generalization Test Conditions

When a child either reached criterion or a total of 250 trials for a concept pair a generalization test was administered during two separate sessions. The generalization test consisted of a total of 60 trials which were spread over two testing times, 24 hours and 48 hours, following acquisition.

Each member of the concept pair was tested with five different exemplars, two novel and three previously seen twice during baseline, on each of three generalization tasks. Generalization task one consisted of items where the irrelevant dimension position varied between members while irrele-

vant dimensions color and form remained constant. Generalization task two had items which varied the irrelevant dimensions position and color between members and form remained constant. Generalization task three consisted of items varied on the irrelevant dimensions position, color and form.

Counterbalancing for position, each of the five exemplars at each generalization task were presented twice, once on the right position and once on the left. Each concept member was tested a total of 30 times, ten trials at each generalization task (one, two and three). A total of 60 trials were required to test both members of a concept pair.

The generalization test series randomly altered the order and position of correct responses for items at each generalization task - one, two and three.

Child responses to generalization test items were not consequated by the experimenter. The instructional procedure consisted of presenting stimulus items representing a concept pair and requesting the child to touch the named concept ("Touch ____"). Only one concept member was requested at one time. This stimulus pair reappeared later in the series where the other member was requested.

A series of competency items, identical to those used during the baseline condition, interspersed through the generalization test, were consequated by the examiner. The rationale for this procedure is identical to that described in the baseline condition.

Individual testing sessions for generalization are

further discussed in Chapter V under the section dealing with generalization results. Appendix D gives a description of items used during the generalization test.

Inter-Observer Reliability

Inter-observer reliability was not taken during the acquisition of concepts in each experimental condition. The nature of the teaching procedures and the lack of subjectivity in deciding a correct or incorrect response did not seem to warrant it.

Inter-observer reliability was taken on fifty percent of total trials for all generalization tests for each child. Four individuals assisted in gathering inter-observer reliability data. Each rater sat off to the left of the child and recorded their responses following each trial. The raters had only to indicate if the child was correct or incorrect following the experimenter's instruction. Following the session each rater and the experimenter computed a reliability score by dividing the number of agreements by the number of agreements plus disagreements and multiplying the resultant number by 100.

There was a total of 16 inter-observer reliability scores computed with a minimum of thirty trials each. Rater one completed eight reliability checks with an average agreement of 100%; rater two completed three checks with an average agreement of 98%; rater three completed two checks with an average agreement of 100%, while rater four completed three checks with an average agreement of 100%. The

slightly lower average agreement score for rater two was the result of a child turning sideways, blocking the view on one trial. The individual agreement scores for rater two were 100%, 96% and 100%.

The next chapter presents the results of group and individual performance in each of the four experimental conditions during acquisition and resulting generalization.

CHAPTER V

RESULTS

The results of the present experiment are reported in two sections. The first deals with concept acquisition, while the second is concerned with concept generalization.

Concept Acquisition

The results of this section are presented first in relation to each experimental hypothesis. Secondly, individual subject data is reported to show a more molecular view of individual variability during acquisition.

The four experimental conditions under investigation are as follows: Condition one, two exemplars with one irrelevant dimension (position) varied between pairs; condition two, four exemplars with one irrelevant dimension (position) varied between pairs; condition three, two exemplars with three irrelevant dimensions (position, color, form) varied between pairs; condition four, four exemplars with three irrelevant dimensions (position, color, form) varied between pairs.

Hypothesis One

Hypothesis one stated that there would be an increase in the number of trials to criterion as the interdimensional

stimulus variation (IER) increases. IER variation refers to the number of irrelevant dimensions which vary between concept pairs during training. There are two levels of IER variation, level one has only one irrelevant dimension changing between concept pairs (position), while level two has three irrelevant dimensions changing between concept pairs (position, color, form).

This overall comparison between IER levels one and two can first be made by taking the average number of trials to criterion for all children over conditions one and two compared with the average number of trials to criterion for conditions three and four.

Table four shows individual child data for total trials to criterion (TTC) in experimental conditions one, two, three and four; total trials to criterion (TTC) for conditions one and two compared to three and four, plus group mean scores for average total trials to criterion (TTC) in each experimental condition and average total trials to criterion (TTC) for conditions one and two and then three and four.

The mean value for conditions one and two TTC is 264, while the mean value for conditions three and four TTC is 508. The mean value for conditions three and four TTC is 1.9 times greater than the mean value for conditions one and three TTC.

Two additional comparisons are made showing the effect of increasing interstimulus variation from one irrelevant to three irrelevant dimensions at each level of number of exem-

Table IV

Total Trials to Criterion for Conditions One and Two Versus Conditions Three and Four

	Condition One	Condition Two	Total One & Two	Condition Three	Condition Four	Total Three & Four
Child						
1	64	64	128	160	160	320
2	96	160	256	128	112	240
3	96	96	192	256	384	640
4	160	320	480	384	448	832
TOTAL	416	640	1056	928	1104	2032
$\bar{m}$	104	160	264	232	276	508

plars. Number of exemplars (NE) refers to the number of examples which are used within a teaching set during training. Level one NE, consists of two exemplars for each member of the concept pairs, while level two NE has four exemplars for each member of the concept pairs.

A comparison of increasing interdimensional stimulus variation (IER) at NE level one is made by comparing the TTC for condition one to three, while at level two NE the comparison is between TTC for conditions two and four. Table four shows this comparison for individual child data for condition one versus condition three and condition two versus condition four in addition to group data for each of these comparisons.

These results show that at level one NE there is an increase in average total trials to criterion for condition three ($\bar{m}$ = 232) over condition one ($\bar{m}$ = 104). Average total trials to criterion (TTC) for condition three are 2.23 times greater than condition one. There is a similar trend observed at level two NE where interdimensional stimulus variation (IER) increases from one irrelevant (condition two) to three irrelevant dimensions (condition four). Condition four ($\bar{m}$ = 276) has 1.73 greater number of TTC than condition two ($\bar{m}$ = 160).

The results from group data are taken as supportive of hypothesis one on the overall comparison and at each level of number of exemplars. That is, as interdimensional stim-

ulus variation increases there is an increase in the total number of trials to criterion.

<u>Individual Summary Results</u>:

Figure two shows the individual child data for total trials to criterion, comparing conditions three and four with conditions one and two. It should be noted that these conditions were not taught simultaneously but are presented together for purposes of visual analysis. Table five shows the order of introduction for condition pairs and concept pairs taught in these conditions for each subject.

As is shown in Table five, experimental condition pairs (condition one with condition two and three with four) were counterbalanced for order of introduction across the four experimental subjects. In addition, each concept pair was represented in each of the experimental conditions. This was done to minimize any effect due to order of introduction for experimental condition pairs and difficulty between individual concept pairs.

The results in Figure two indicate a replication of greater number of trials to criterion for condition three over condition one across all children. The comparison between conditions two and four shows a greater number of trials to criterion for condition four over two replicated across three children. Child two showed the reverse with greater number of trials to criterion for condition two over condition four.

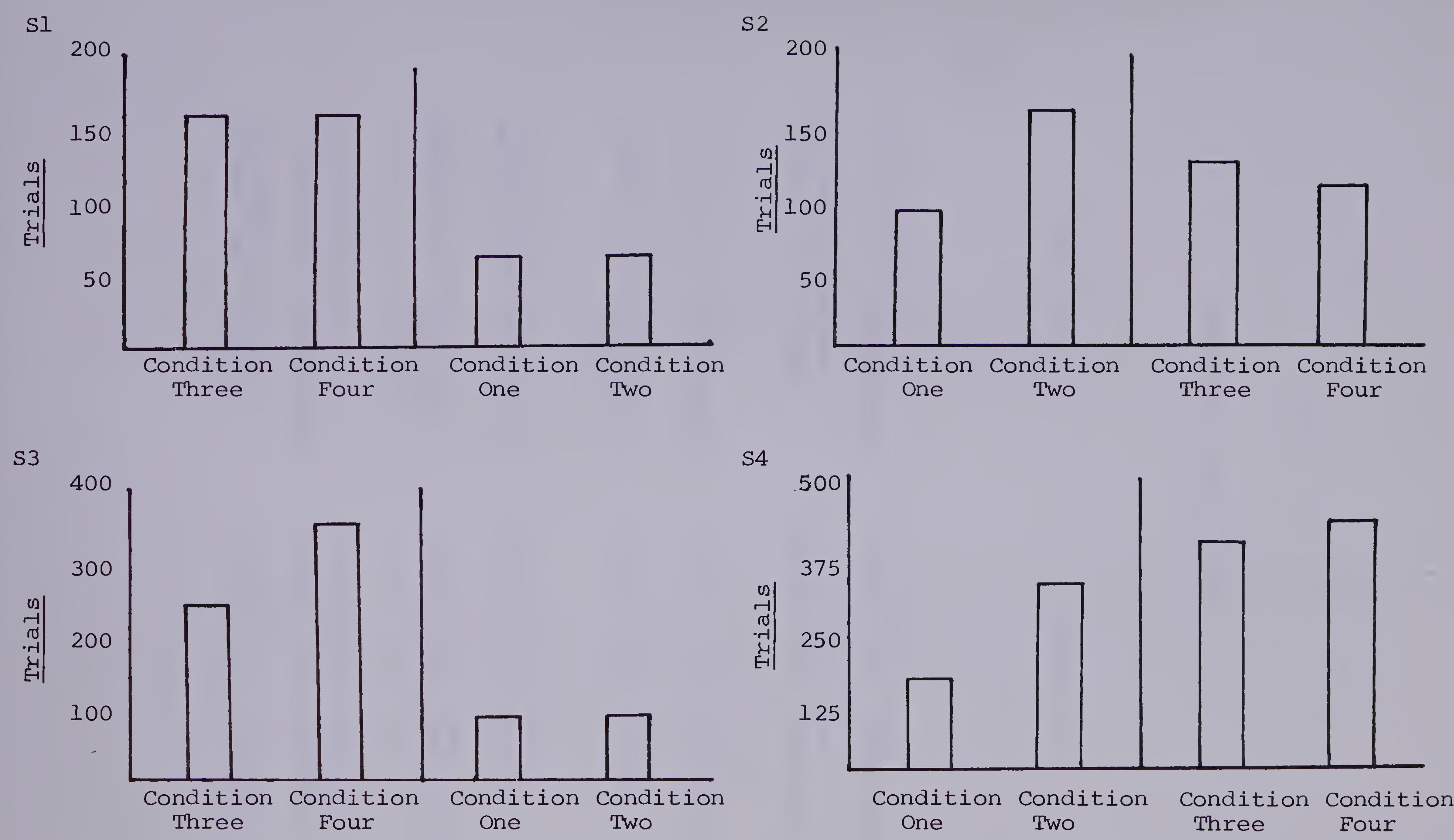


Figure 2: Comparison of total trials to criterion comparing condition one with condition two and condition three with condition four for each subject.

Table V

Order of Introduction

Subject	First Pair	Second Pair
1	a) Condition Four Concept pair Empty-Full	a) Condition One Concept pair Straight-Curved
	b) Condition Three Concept pair In-Out	b) Condition Two Concept pair Long-Short
2	a) Condition One Concept pair Long-Short ,	a) Condition Three Concept pair Empty-Full
	b) Condition Two Concept pair Straight-Curved	b) Condition Four Concept pair In-Out
3	a) Condition Four Concept pair Long-Short	a) Condition One Concept pair In-Out
	b) Condition Three Concept pair Straight-Curved	b) Condition Two Concept pair Empty-Full
4	a) Condition One Concept pair Empty-Full	a) Condition Three Concept pair Long-Short
	b) Condition Two Concept pair In-Out	b) Condition Four Concept pair Straight-Curved

When averaging across children the reverse effect for subject two when comparing condition two versus four is lost. These individual results strongly support the experimental hypothesis at level one NE and partially support it at level two NE.

Hypothesis Two

Hypothesis two stated that there would be an increase in the number of trials to criterion as the number of exemplars (NE) increases. Number of exemplars has two levels of variation, level one has two exemplars of a concept class within the teaching set, while level two has four.

In a similar format to hypothesis one the results for hypothesis two will be presented as an overall comparison for two versus four exemplars and then at each level of IER variation. The overall comparison for number of exemplars level one compared with level two is made by taking the average number of trials to criterion (TTC) for all subjects over conditions one and three compared to conditions two and four.

Table six shows individual subject and group data for each experimental condition across conditions one and three and conditions two and four totals.

The mean value for conditions one and three TTC is 336 while the TTC for conditions two and four totals is 435. The TTC for conditions two and four is 1.3 times greater than TTC for conditions one and three.

Table VI

Total Trials to Criterion for Conditions One and Three Versus Conditions Two and Four

	Condition One	Condition Three	Total One & Three	Condition Two	Condition Four	Total Two & Four
Child						
1	64	160	224	64	160	224
2	96	128	224	160	112	272
3	96	256	352	96	384	480
4	160	384	544	320	448	768
TOTAL	416	928	1344	640	1104	1744
$\bar{m}$	104	232	336	160	276	436

Each level of number of exemplars is further compared at the two levels of IER variation by average total trials to criterion for condition one compared to condition two and for condition three with four. Table four includes the comparison of increased number of exemplars at two levels of IER variation. This comparison at level one IER variation is seen between conditions one and two while at level two IER variation it is between conditions three and four.

There is a similar trend in average total trials to criterion when comparing two versus four exemplars at the two levels of IER variation. With one irrelevant dimension, condition two ($\bar{m}$ = 160) has 1.54 greater TTC than condition one ($\bar{m}$ = 104); with three irrelevant dimensions condition four ($\bar{m}$ = 276) has 1.20 greater TTC than condition three ($\bar{m}$ = 232).

The increased trials to criterion as the number of exemplars increases first for the overall comparison and secondly, at each level of IER variation are seen as supportive of hypothesis two.

Individual Summary Results:

Figure three shows each child's data for total trials to criterion comparing conditions one with two and three with four. These conditions were taught simultaneously and the results are presented graphically in the order of pair introduction during training.

The results for comparing these conditions shows the strong effect of IER variation. However, the comparison of

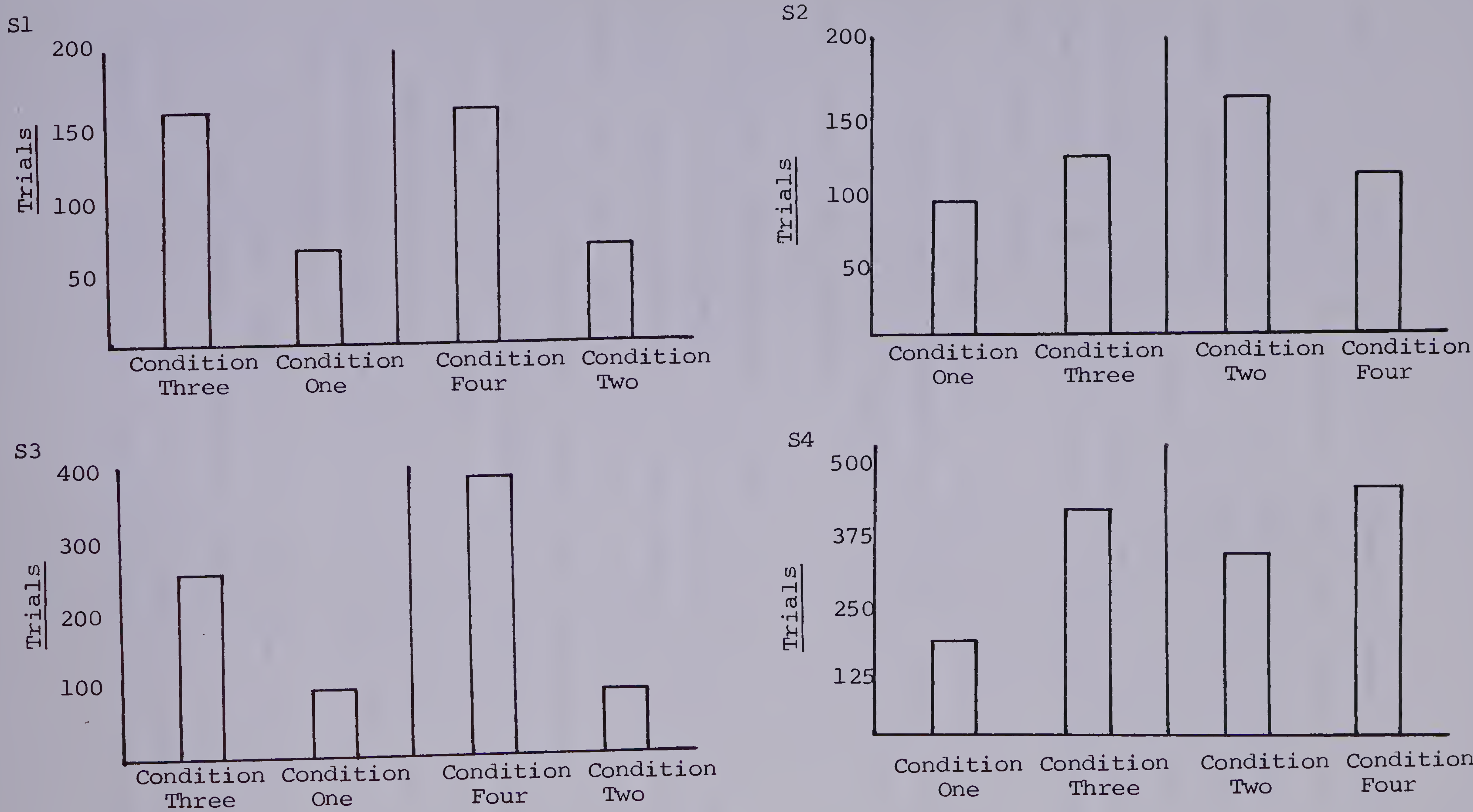


Figure 3: Comparison of total trials to criterion comparing conditions one with three and two with four for each subject.

condition one with condition two showed a greater number of trials to criterion for condition two over condition one for children two and four, while children one and three showed no difference. The comparison of condition three with four reveals a greater number of trials to criterion for condition four over three for children three and four while there was no difference for child one, and child two showed the reverse effect.

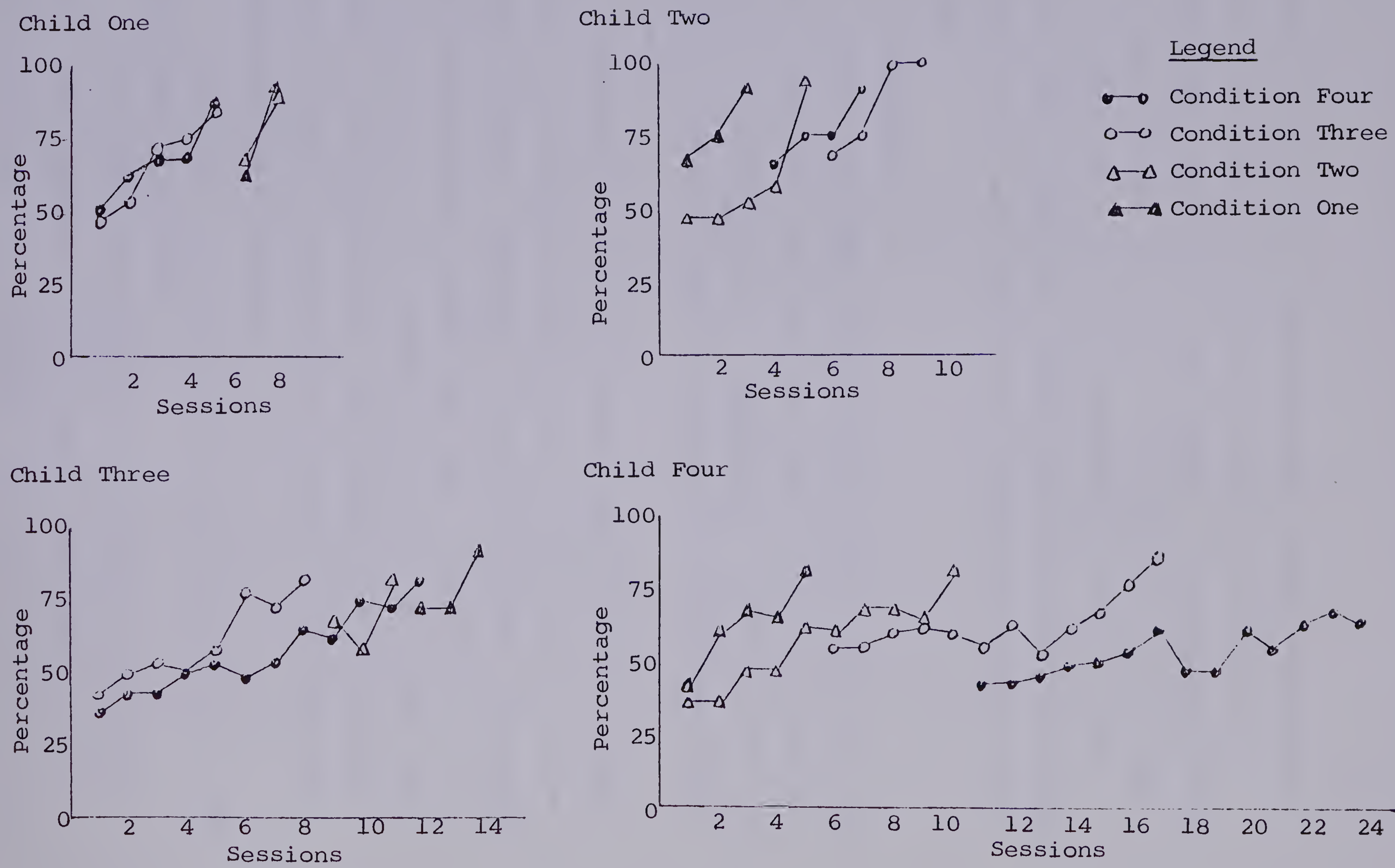
Averaging results across children as was previously done obscures these individual variations. These individual results only partially support hypothesis two at levels one and two IER variation.

Individual Child Acquisition Data

The section reports a more molecular view of an individual child's acquisition of each member of a concept pair within each experimental condition. Figure four shows percentage correct scores for concept pairs each session. These scores are collapsed across both members of a concept pair taught in the condition. Figure four portrays an overall comparative view for all children showing the order of introduction for experimental conditions and number of sessions to criterion.

The next four figures, (five, six, seven and eight) show individual subject acquisition results for each member of the concept pairs taught in their respective experimental conditions. These figures have the results for each session

Figure 4: Percentage correct each session



plotted in four trial blocks for each concept pair member. There was a total of thirty-two trials per session for a concept pair, sixteen for each member. A percentage correct score out of four was computed for each member of the pair and plotted over time. This presentation format allows a closer look at individual variability both within and between conditions for each member of the concept pair during acquisition.

Child One:

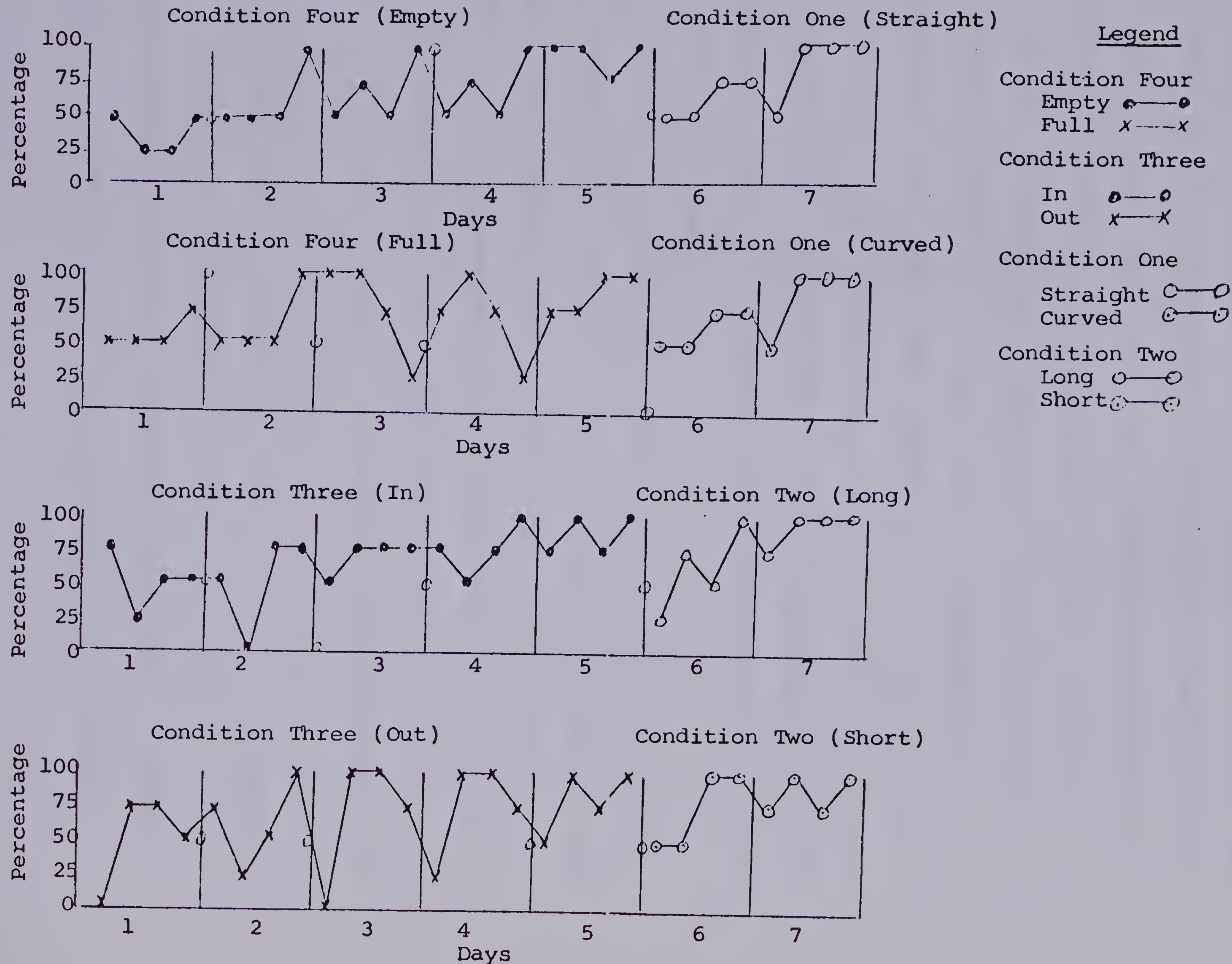
Figure five shows the individual data for child one. The graph shows that conditions three and four were introduced first and the concept pairs "empty-full" (condition four) and "in-out" (condition three) were acquired in the same number of sessions.

In condition four, the concept "full" appears to have greater variability of scores around chance level than the concept "empty". Similarly, in condition three, the concept "out" appears more variable than "in". Conditions one and two show similar trends between conditions and within each member of the concept pair. The variability with concept pairs and between conditions one and two is less than for conditions three and four.

Concepts in conditions one and two were acquired in the same number of sessions.

Daily probes were taken for the concepts to be taught next. Four probes were taken, two for each member of the next concept pair to be trained. The probe items were selec-

Figure 5: Child One



ted from the range of stimulus cards to be included in the training set. Probe items were interchanged at random from day to day and were always at an IER level three variation.

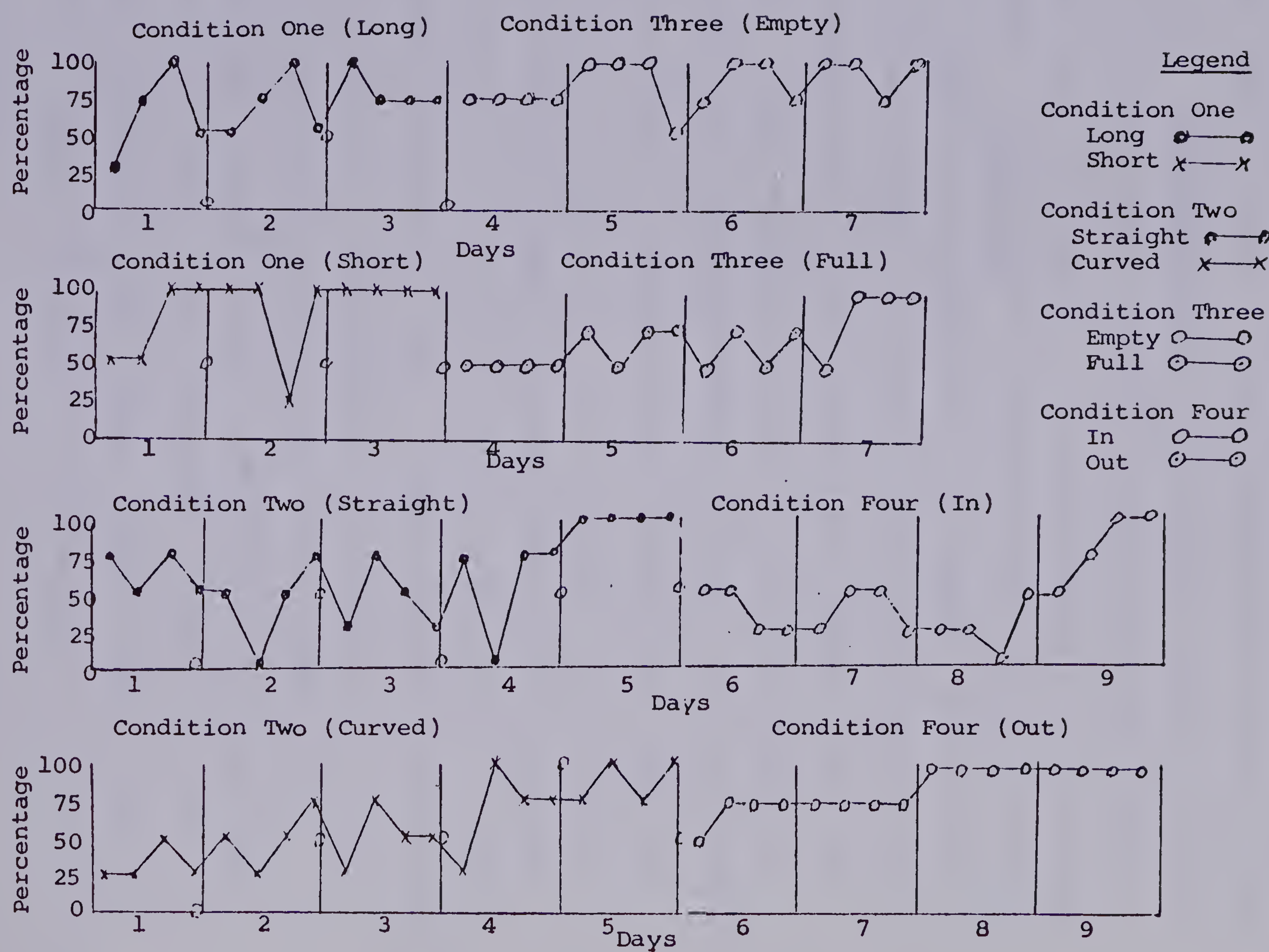
A percentage score of 0, 50, 100% was calculated and plotted over time by an open circle. Child one scored a total of 17% and 50% for "straight" and "curved" probe items during training of "empty-full" in condition four. Probe scores for "long" and "short" were 30% and 50% while "in" and "out" were trained in condition three. This probe data suggests that the concept pairs "straight-curved" and "long-short" were not acquired prior to instruction in their respective conditions.

There is no evidence of overtraining where one member of a concept pair reaches criterion before the other. Overtraining can possibly occur as the criterion for mastery was 80% for both members of the concept pair. In a situation of overtraining one member will reach a criterion of 80% while the other remains below 80%. Therefore while one member continues to require training to criterion, the other receives additional training or practise. The incidence of overtraining in the present data and its relationship to acquisition and generalization under the present instructional procedures will be discussed in the next chapter.

Child Two:

Figure six shows the individual data for child two. Conditions one and two were introduced first with condition

Figure 6: Child Two



one being acquired first. This effect resulted in a staggered introduction for conditions three and four.

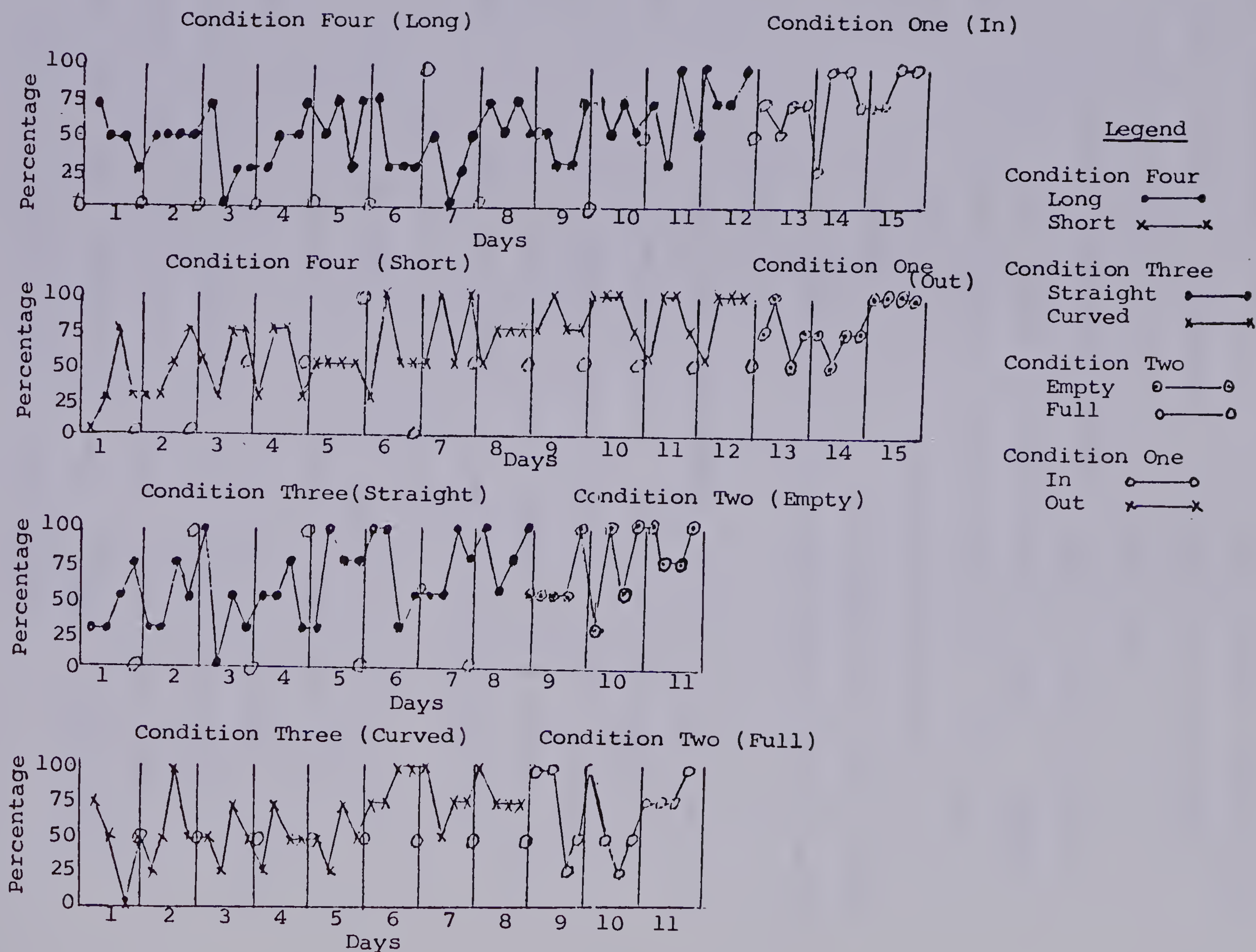
In condition one, the concept "long" appears more variable than "short" while in condition two, both members of the concept pair "straight-curved" show similar degrees of variability in performance. Condition three was taught for two sessions with condition two. The variability of the concept pair members "empty-full" in condition three appears similar, however, "empty" was consistently at a higher level of percentage correct scores. Condition four presents an interesting situation of overtraining for one member of the concept pair. "Out" reached individual criterion during day eight, while "in" required another session. Prior to session nine the concept "in" appears more variable than "out" and consistently at a lower level of accuracy. During session nine, we observe a gradual increase from 50 to 100 percent correct for "in" over the four blocks, while "out" is being overtrained. Generally, "out" is less variable is performance and the situation of overtraining may have assisted in discrimination and acquisition of the concept member "in".

Probe data for child two, although showing a different variability of results when compared to child one, does indicate that the concept pairs "empty" (33%) - "full" (33%) and "in" (30%) - "out" (50%) were not learned prior to instruction.

Child Three:

Figure seven shows the individual data for child three.

Figure 7: Child Three



Conditions three and four were introduced first with condition three being acquired first. This effect resulted in a similar trend of staggered introduction for conditions one and two as was the case for subject two.

In condition four the concept "<u>long</u>" appears slightly less variable than "short" for days one to eight. "<u>Short</u>" reached individual criterion during day eight with overlearning during day ten where the individual criterion was again met. Individual criterion for "<u>short</u>" fell off during day eleven and was regained on day twelve where criterion was met for both members of the pair.

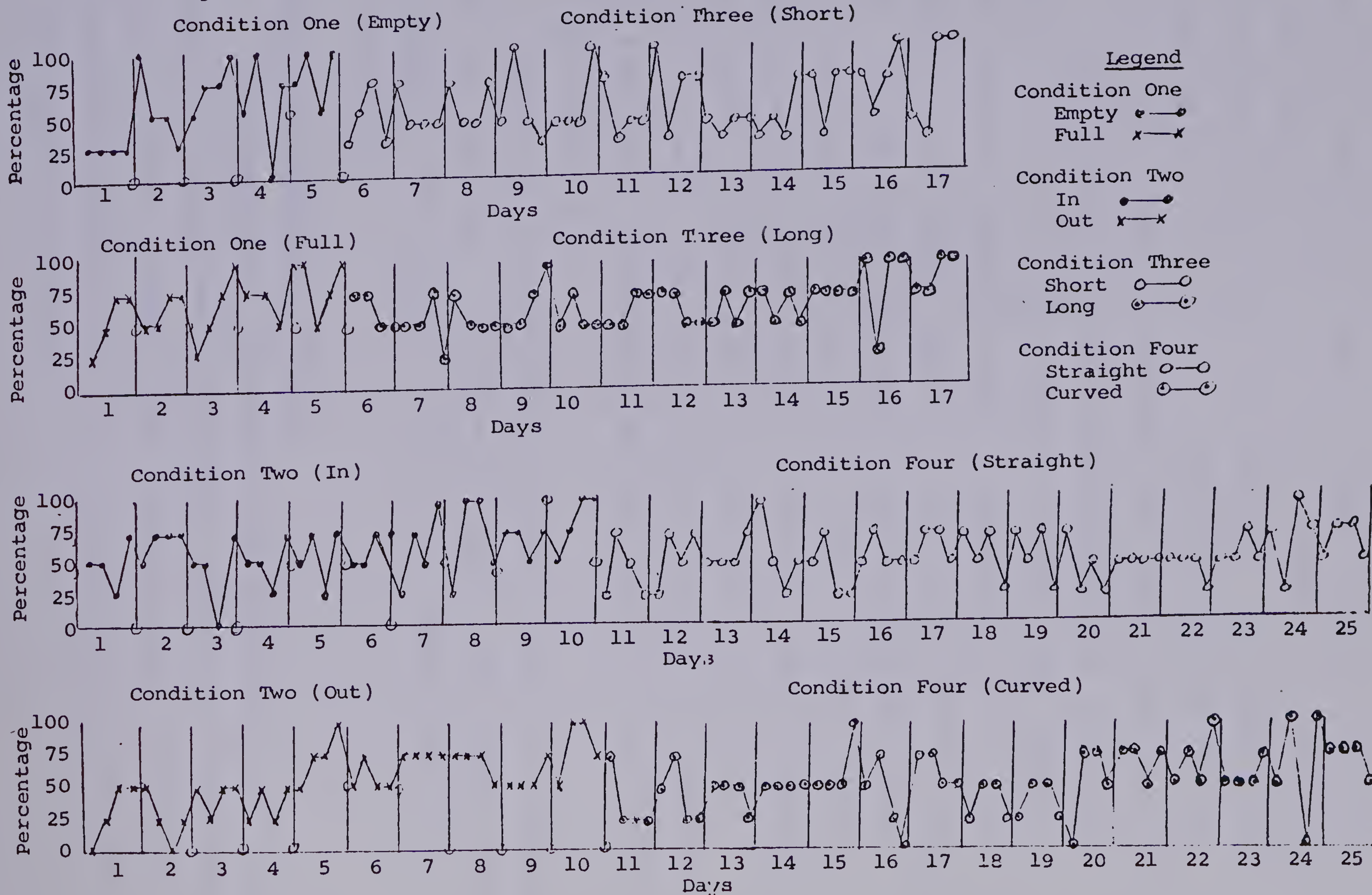
The variability for each member of the concept pair in condition three appears similar. There is one example of overlearning for "<u>curved</u>" during day six. This phenomenon dropped off in the next session, however, in a similar manner to condition four, criterion for both concept pairs was attained within the next two sessions.

Concept pair members in both conditions one and two show a similar degree of variability. Both conditions one and two were trained to criterion in the same number of days with no evidence of overtraining for any concept member. Similarly, with children one and two, probe data shows that concept pairs "<u>in</u>" (25%)-"<u>out</u>" (42%), "<u>empty</u>" (38%)-"<u>full</u>" (50%) were not learned prior to instruction.

<u>Child Four</u>:

Figure eight shows the individual data for child four.

Figure 8: Child Four



Conditions one and two were introduced first and condition one was acquired first. This effect resulted in a staggered introduction for the next two conditions. The necessity of staggering the introduction of conditions from the second pair has been previously observed with children two and three.

The variability between concept members in conditions one and two and betweeen conditions appears very similar. The results are generally quite variable even with the exception of the concept "<u>out</u>" in condition two from the seventh day. At this point "<u>out</u>" appears less variable than "<u>in</u>".

Condition three results show a similar degree of variability between concept members as was observed in conditions one and two. Condition four did not reach criterion for mastery; however, the criterion for termination of training was a maximum to two hundred and forty trials for each member of the concept pair.

Probe data indicates similar trends as shown for children one, two and three. Concept pairs "<u>short</u>" (10%)-"<u>long</u>" (50%), "<u>straight</u>" (35%)-"<u>curved</u>" (15%) were not learned prior to instruction.

<u>Summary of Acquisition Results</u>

<u>Hypothesis One</u>: The results from the group data strongly support this experimental hypothesis. That is, as interdimensional stimulus variation (IER) increased from one to three irrelevant dimensions there was also an increase in average total trials to criterion.

This effect was observed for group data at each level of number of examples (NE). As interdimensional stimulus variation increased from one to three irrelevant dimensions at level one NE (two exemplars) and two NE (four exemplars) there was an increase in average total trials to criterion.

Individual child results show a replication of this effect across all children at level one NE while only three children show the effect at level two NE. Individual child replication strongly supports the hypothesis of increased total trials to criterion as interstimulus variation increased from one to three irrelevant dimensions at level one NE, while only partially supporting the hypothesis at level two NE.

Hypothesis Two: The results from group data also strongly support this experimental hypothesis. Support for the experimental hypothesis was observed both for the overall comparison of increased trials to criterion as number of exemplars increased and as number of exemplars increased at each level of interstimulus variation (IER).

That is, as number of exemplars increased from level one (two exemplars) to level two (four exemplars) first at level one IER (one irrelevant dimension) and secondly at level two IER (three irrelevant dimensions) there was an increase in the average total trials to criterion. Individual child replication only weakly supports the hypothesis for increased trials to criterion as number of exemplars increased. This support is weak due to the fact that two subjects showed an effect for increased trials to criterion as number of exemplars in-

creased at both levels of IER variation, while two did not.

Generalization Results

A generalization test was administered in two parts, twenty-four hours (next session) and forty-eight hours (second session) following the attainment of criterion for acquisition. Each generalization test was administered in the same manner. Due to the fact that responses on the generalization items were not consequated by the experimenter, a series of competency items were randomly mixed into the test. Competency items were those items each child identified correctly during the pre-baseline tests. These items had no relationship to the concepts being tested. Child responses to the competency items were consequated by the experimenter. This procedure established an intermittent schedule of feedback to each child during the generalization tests. This was done to ensure a greater probability of on-task and response behavior by each child. The response topography ("touch ___") for the competency items was the same as for generalization test items. No attempt was made to demonstrate the effectiveness of this procedure, however, it has been previously used in a similar manner by Panyan and Hall (1978).

The generalization test consisted of fifteen novel exemplars for each member of a concept pair. To ensure that a position effect was not in effect a correct response to each item occurred in both the right and left position. The generalization test was broken into three separate components.

The first component, generalization task one (GT1), consisted of five exemplars where item pairs varied on only one irrelevant dimension (position). The second component, generalization task two (GT2), consisted of five exemplars where item pairs varied in two irrelevant dimensions (position and color). The third component, generalization task three (GT3), consisted of five exemplars where item pairs varied in three irrelevant dimensions (position, color and form). Items from the three components were presented in a random order during the generalization test. Thirty trials were administered on day one of the generalization test and the same thirty trials in a different order were again presented on day two. Table seven shows the number of presentations for each concept in a pair for each component task and totals for each day of the generalization test.

Results of the generalization test will be reported in raw scores and percentage correct scores.

Hypothesis Three

Hypothesis three stated that there would be greater number of total generalization responses observed as the number of exemplars (NE variation) increases from level one (two exemplars) to level two (four exemplars). The overall comparison for this hypothesis is made by taking the average total trials correct on generalization for all children for conditions one and three compared to conditions two and four. The total generalization scores for each subject were calcu-

Table VII

Generalization Test

	Day One (Number of Exemplars)		Total	Day Two (Number of Exemplars)		Total	Totals for Day One and Two (Number of Exemplars)		Total Number of Exemplars
	Concept Pair			Concept Pair			Concept Pair		
	A	B		A	B		A	B	
Components									
Task One	5	5	10	5	5	10	10	10	20
Task Two	5	5	10	5	5	10	10	10	20
Task Three	5	5	10	5	5	10	10	10	20
TOTAL			30			30			60

lated by totalling their respective scores at each of the three generalization tasks (GT1 + GT2 + GT3).

Table eight shows individual child total correct and overall average total correct for conditions one and three compared to conditions two and four.

The results in Table eight show an overall average total of $\bar{m}$ = 84 (70%) generalization responses for conditions one and three while the average total for conditions two and four is $\bar{m}$ = 82 (69%). This difference of two responses or one percentage point does not appear great enough to be supportive of hypothesis three.

Two additional comparisons can be made isolating the effects of increasing the number of exemplars from two to four at each level of IER variation. This comparison is made by comparing the average total generalization responses for condition one with condition two and condition three with condition four. Comparison of condition one with two looks at increasing the number of exemplars from two to four with one irrelevant dimension varied, while condition three with four investigates the increase in exemplars with three irrelevant dimensions varied.

The results from Table eight show the average total correct generalization responses for condition two ($\bar{m}$ = 42) to be five responses less than condition one ($\bar{m}$ = 38). This raw score difference represents seven percentage points, and although total generalization to condition one is greater than condition two, this difference does not appear large

Table VIII

Generalization Scores and Percentages Comparing Total of Conditions One and Three Versus Two and Four

	Condition One Total Scores		Condition Three Total Scores		Total One & Three Scores		Condition Two Total Scores		Condition Four Total Scores		Total Three & Four Scores	
Child	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
1	46	76	57	95	103	85	48	80	60	100	108	90
2	43	72	46	77	89	74	39	65	44	73	83	69
3	44	73	38	63	82	68	41	68	40	67	81	68
4	33	55	31	52	62	52	25	42	30	50	55	46
TOTAL	166		172		336		153		174		327	
$\bar{m}$	42	70	43	72	84	70	38	63	44	73	82	69

enough to support the hypothesis at level one IER variation. The comparison between conditions three ($\bar{m}$ = 43) and four ($\bar{m}$ = 44) shows a small difference in average total generalization responses and is seen as not supportive of the hypothesis at level two IER variation.

Hypothesis Four

Hypothesis four stated that as the number of exemplars increased from two to four there would be a difference in the total generalization responses on each on the component tasks in the generalization test.

Generalization Task One (GT1)

Table nine shows the results of this comparison across conditions one and three totals compared with conditions two and four totals for GT1. Average total generalization to GT1 is slightly greater for conditions one and three ($\bar{m}$ = 33) than for the same total in conditions two and four ($\bar{m}$ = 30).

Comparing average total generalization to GT1 at each level of IER variation shows that condition one is slightly higher than condition two and condition three is similarly higher than condition four. These results do not support the experimental hypothesis at GT1.

Generalization Task Two (GT2)

Table ten shows the comparisons of average total correct generalization responses for condition one and three with two and four. There is no difference between the total for conditions one and three and two and four at GT2. The additional

Table IX

Total Generalization Responses to Generalization Task One (G.T.1)

	Condition One Scores		Condition Three Scores		Condition One & Three Scores		Condition Two Scores		Condition Four Scores		Condition Two and Four Scores	
Child	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
1	20	100	19	95	39	98	20	100	20	100	40	100
2	20	100	17	85	37	93	16	80	14	70	30	75
3	18	90	11	55	29	73	17	85	13	65	30	75
4	15	75	12	60	27	68	12	55	9	45	21	53
TOTAL	73		59		132		65		56		121	
$\bar{m}$	18	90	15	75	33	83	16	80	14	70	30	75

Table X

<u>Total Generalization Responses to Generalization Task Two (GT2)</u>

	Condition One Scores		Condition Three Scores		Condition One & Three Scores		Condition Two Scores		Condition Four Scores		Condition Two & Four Scores	
	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
Child												
1	17	85	20	100	37	93	19	95	20	100	39	98
2	16	80	15	75	31	78	14	70	14	70	28	70
3	18	90	11	55	29	73	16	80	13	65	29	73
4	10	50	8	40	18	45	9	45	11	55	20	50
TOTAL	61		54		115		58		58		116	
$\bar{m}$	15.3	77	13.5	68	29	73	15	75	15	75	29	73

comparisons at each level of IER variation shows no difference at level one (condition one compared to condition two and a slight difference at level two favoring condition four ($\bar{m}$ = 15) over three ($\bar{m}$ = 13.5). These results do not support the experimental hypothesis at GT2.

Generalization Task Three (GT3)

Table eleven shows the comparison of average total generalization responses for conditions one and three with two and four. There is no real difference with either the overall comparison for conditions one and three compared to conditions two and four totals at GT3 or with level of IER variation, condition one compared to two and three with four.

These results do not support the experimental hypothesis at GT3.

Individual Generalization Results

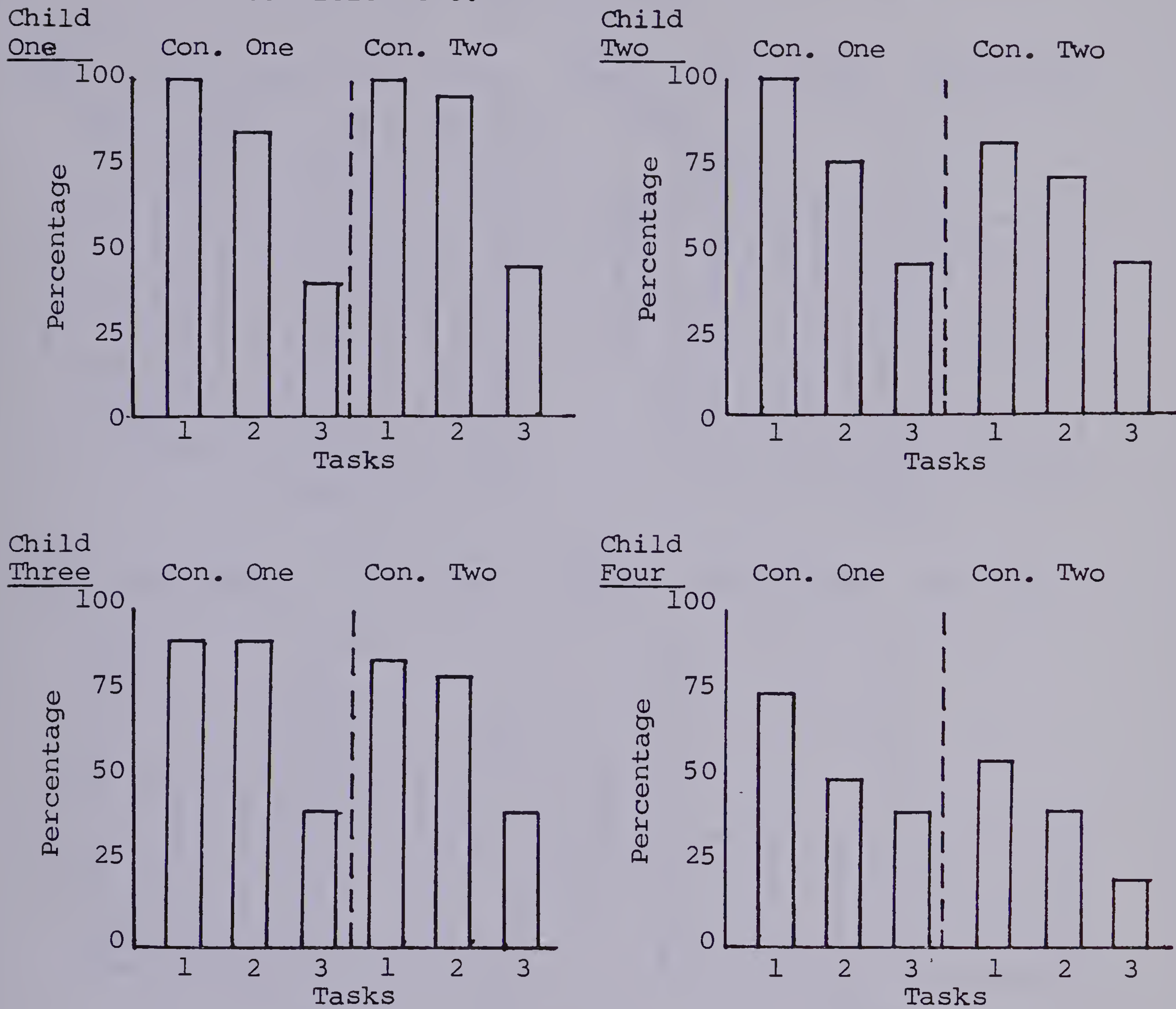
Figure nine shows a comparison of conditions one and two for each task of the generalization test for each child and average totals for all children. The results show visually that conditions one and two appear as a replication of themselves with no real differences within children over the two conditions. Figure ten shows the same comparison for conditions three versus four. These results also show no real difference for subjects between conditions three and four on each component task of generalization.

Table XI

Total Generalization Responses to Generalization Task Three (GT3)

	Condition One Scores		Condition Three Scores		Condition One & Three Scores		Condition Two Scores		Condition Four Scores		Condition Two & Four Scores	
	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
Child												
1	9	45	18	90	27	68	9	45	20	100	29	73
2	8	40	14	70	22	55	8	40	16	80	24	60
3	8	40	16	80	24	60	8	40	15	75	23	58
4	8	40	11	55	19	48	8	40	10	50	18	45
TOTAL	33		59		92		33		61		94	
$\bar{m}$	8.3	42	15	75	23	58	8.3	42	15.3	77	24	60

Figure 9: Percentage of Generalization to tasks one, two and three. Comparison of condition one with condition two.



Average Percentage of Generalization to Tasks Two, Three for Conditions One, Two

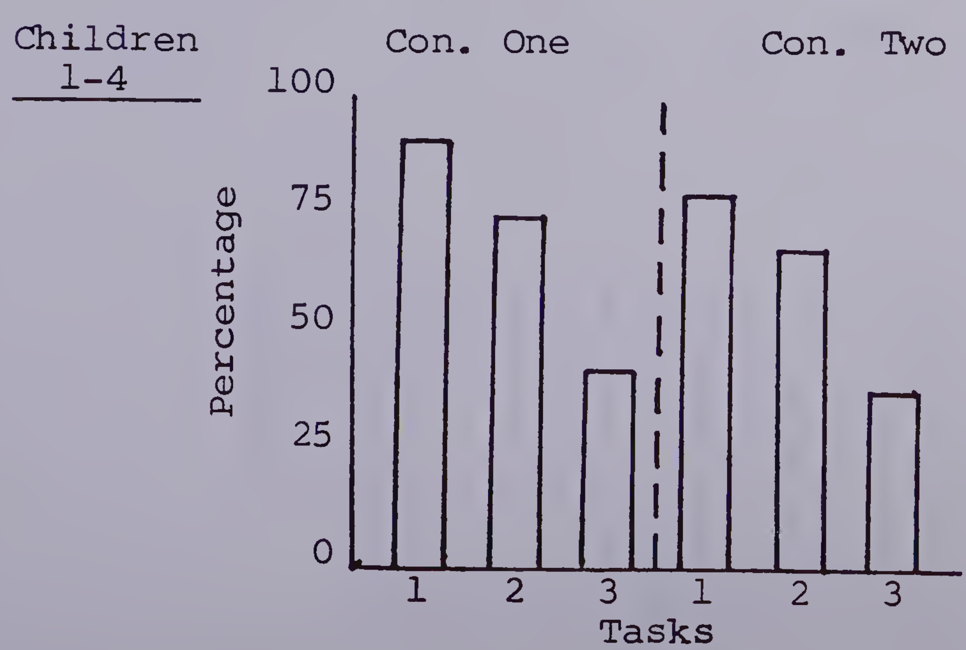
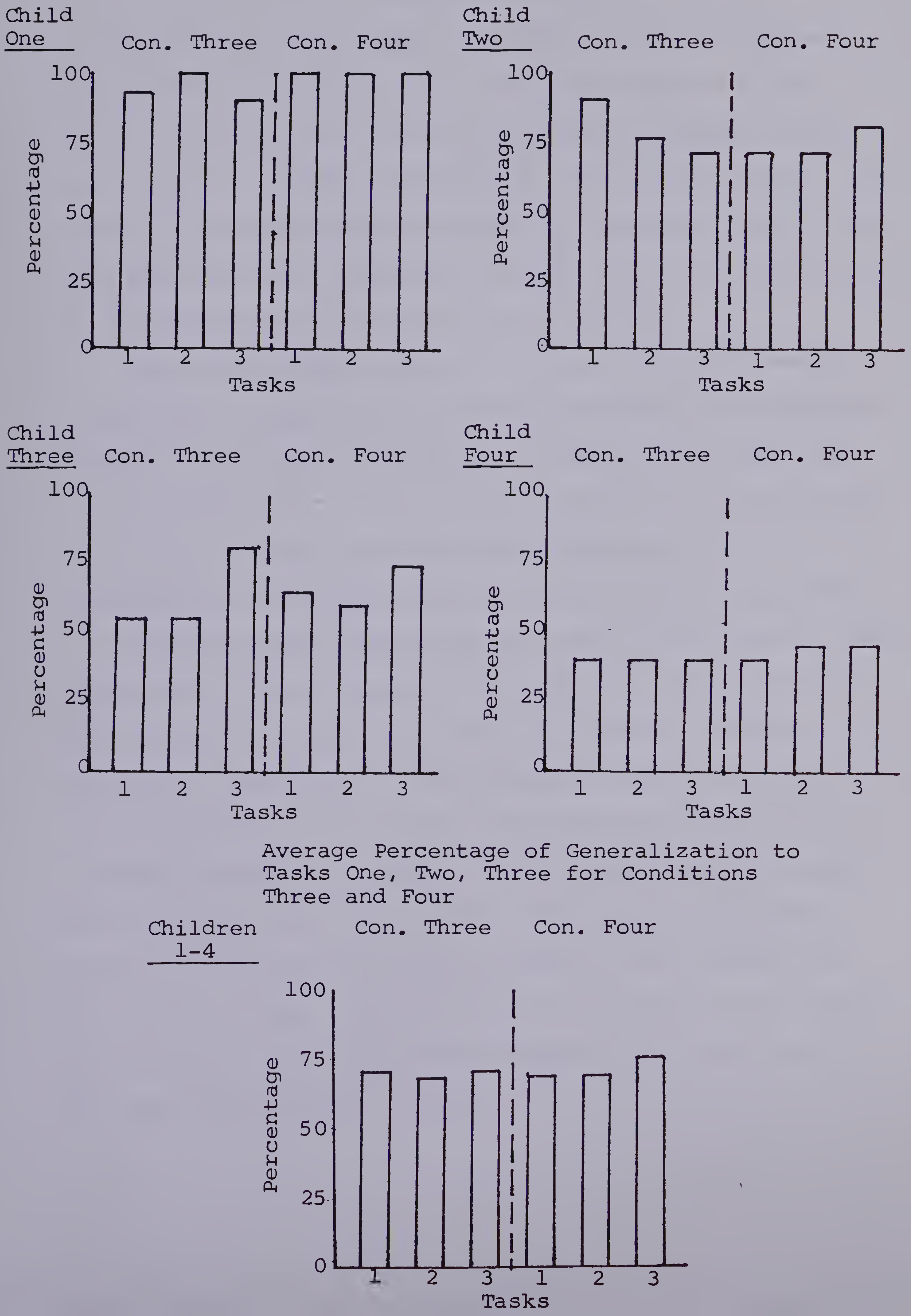


Figure 10: Percentage of generalization to tasks one, two and three. Comparison of condition three with condition four.



Hypothesis Five

Hypothesis five stated that there would be a greater total number of generalization responses observed as the level of interstimulus variation increases from one irrelevant to three irrelevant dimensions. The initial overall comparison for this hypothesis is made by comparing the average total generalization responses for all children for conditions one and two versus conditions three and four.

Table twelve shows individual child total and overall average total correct generalization responses for this comparison. The overall comparison in Table twelve does show a slight difference between the average total for conditions one and two ($\bar{m}$ = 80) and conditions three and four ($\bar{m}$ = 87). This comparison does not support the experimental hypothesis.

Two additional comparisons are made for increasing interstimulus variation at each level of NE variation. Increasing interstimulus variation from one to three irrelevant dimensions at level one NE (two exemplars) shows no difference between condition one ($\bar{m}$ = 42) and condition three ($\bar{m}$ = 43). A similar comparison at level two NE (four exemplars) shows that condition four ($\bar{m}$ = 44; $\bar{m}$% = 73%) is ten percentage points and six raw score points greater than condition two ($\bar{m}$ = 38; $\bar{m}$% = 63%). These results do not support the hypothesis at level one NE and weakly support it at level two NE for total generalization scores.

Table XII

Generalization Scores and Percentages Comparing Total of Conditions One and Two Versus Three and Four

	Condition One Scores		Condition Two Scores		Total One and Two Scores		Condition Three Scores		Condition Four Scores		Total Three & Four Scores	
	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
Child												
1	46	76	48	80	94	78	57	95	60	100	117	98
2	43	72	39	65	82	68	46	77	44	73	90	75
3	44	73	41	68	85	71	38	63	40	67	78	65
4	33	55	25	42	58	48	31	52	30	50	61	51
TOTAL	166		153		319		172		174		346	
$\bar{m}$	42	70	38	63	80	66	43	72	44	73	87	73

Hypothesis Six

Hypothesis six stated that as interstimulus variation increased from one irrelevant to three irrelevant dimensions there would be a difference in the total generalization responses to each task in the generalization test.

Generalization Task One (GT1)

Table thirteen shows the results of this overall comparison for conditions one and two average total compared to conditions three and four. The difference of fifteen percentage points for conditions one and two average total (88%) over the average total for conditions three and four (73%) would not appear to support the hypothesis at GT1.

Additional comparisons at each level of NE variation show results in opposite direction. At level one NE condition one ($\bar{m}$ = 18; $\bar{m}$% = 90) has a greater number of generalization responses to GT1 than condition three ($\bar{m}$ =15; $\bar{m}$% = 75%), while at level two NE a similar effect is noticed as condition two ($\bar{m}$ = 16; $\bar{m}$% = 80%) is larger than condition four ($\bar{m}$ = 14, $\bar{m}$% = 70%). These results do not support the hypothesis that there will be an increase in generalization responses to GT1 as interstimulus variation increases at each level of NE variation.

Generalization Task Two (GT2)

Table fourteen shows the overall comparison of average total generalization responses for conditions one and two compared to three and four. The results of this comparison

Table XIII

Total Generalization Responses to Generalization Task One (GT1)

	Condition One Scores		Condition Two Scores		Condition One & Two Scores		Condition Three Scores		Condition Four Scores		Condition Three and Four Scores	
	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
Child												
1	20	100	20	100	40	100	19	95	20	100	39	98
2	20	100	16	80	36	90	17	85	14	70	31	78
3	18	90	17	85	35	88	11	55	13	65	24	60
4	15	75	12	55	27	68	12	60	9	45	21	53
TOTAL	73		65		138		59		56		115	
$\bar{m}$	18	90	16	80	35	88	15	75	14	70	29	73

Table XIV

Total Generalization Responses to Generalization Task Two (GT2)

	Condition One Scores		Condition Two Scores		Condition One & Two Scores		Condition Three Scores		Condition Four Scores		Condition Three and Four Scores	
	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
Child												
1	17	85	19	95	36	90	20	100	20	100	40	100
2	16	80	14	70	30	75	15	75	14	70	29	73
3	18	90	16	80	34	85	11	55	13	65	24	60
4	10	50	9	45	19	48	8	40	11	55	19	48
TOTAL	61		58		119		54		58		112	
$\bar{m}$	15.3	77	15	75	29.8	75	13.5	68	15	75	28	70

do not show a large enough difference to support the experimental hypothesis.

The further comparison at level one NE variation shows condition one ($\bar{m}$ = 15.3; $\bar{m}$% = 77%) to be slightly greater than condition three ($\bar{m}$ = 13.5; $\bar{m}$% = 68%), while at level two NE condition two ($\bar{m}$ = 15; $\bar{m}$% = 75%) and four ($\bar{m}$ = 15; $\bar{m}$% = 75%) have identical results. These results do not support the experimental hypothesis at GT2.

Generalization Task Three (GT3)

Table fifteen shows the overall comparison of average total correct generalization responses and percent correct for conditions one and two compared to three and four. This comparison shows a large difference of 14.5 raw score and thirty-six percentage points for conditions three and four total ($\bar{m}$ = 30; $\bar{m}$% = 75%) over conditions one and two ($\bar{m}$ = 15.5; $\bar{m}$% = 39%). These results strongly support the experimental hypothesis at GT3.

Further comparisons at each level of NE variation show support for the hypothesis at both levels one and two. Condition three average percentage score ($\bar{m}$ = 15; $\bar{m}$% = 75%) is thirty-two percent greater than condition one average total ($\bar{m}$ = 8.3; $\bar{m}$% = 42%), while condition four average total ($\bar{m}$ = 15.3; $\bar{m}$% = 77%) is forty percent greater than condition two (m = 7.3; m% = 37%). These results strongly support the experimental hypothesis at GT3.

Table XV

Total Generalization Responses to Generalization Task Three (GT3)

	Condition One Scores		Condition Two Scores		Condition One & Two Scores		Condition Three Scores		Condition Four Scores		Condition Three & Four Scores	
	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%	Raw	%
Child												
1	9	45	9	45	18	45	18	90	20	100	38	95
2	8	40	8	40	16	40	14	70	16	80	30	75
3	8	40	8	40	16	40	16	80	15	75	31	78
4	8	40	4	20	12	30	11	55	10	50	21	53
TOTAL	33		19		62		59		61		120	
$\bar{m}$	8.3	42	7.3	37	15.5	39	15	75	15.3	77	30	75

<u>Individual Generalization Results</u>

Figure eleven shows a comparison of conditions one and three for each task of generalization tested for each child and as average totals for all children.

Children one and two show a similar trend in their results for generalization to GT1 and GT2 between conditions one and three. Generalization to GT3 shows a similar trend between subjects in favor of condition three.

Child three does not replicate the generalization to GT1 and GT2 across conditions one and three which was observed for children one and two. There is greater generalization to GT1 and GT2 for conditions one over three. Generalization to GT3 in favor of condition three is seen as a replication of the same effect for children one and two.

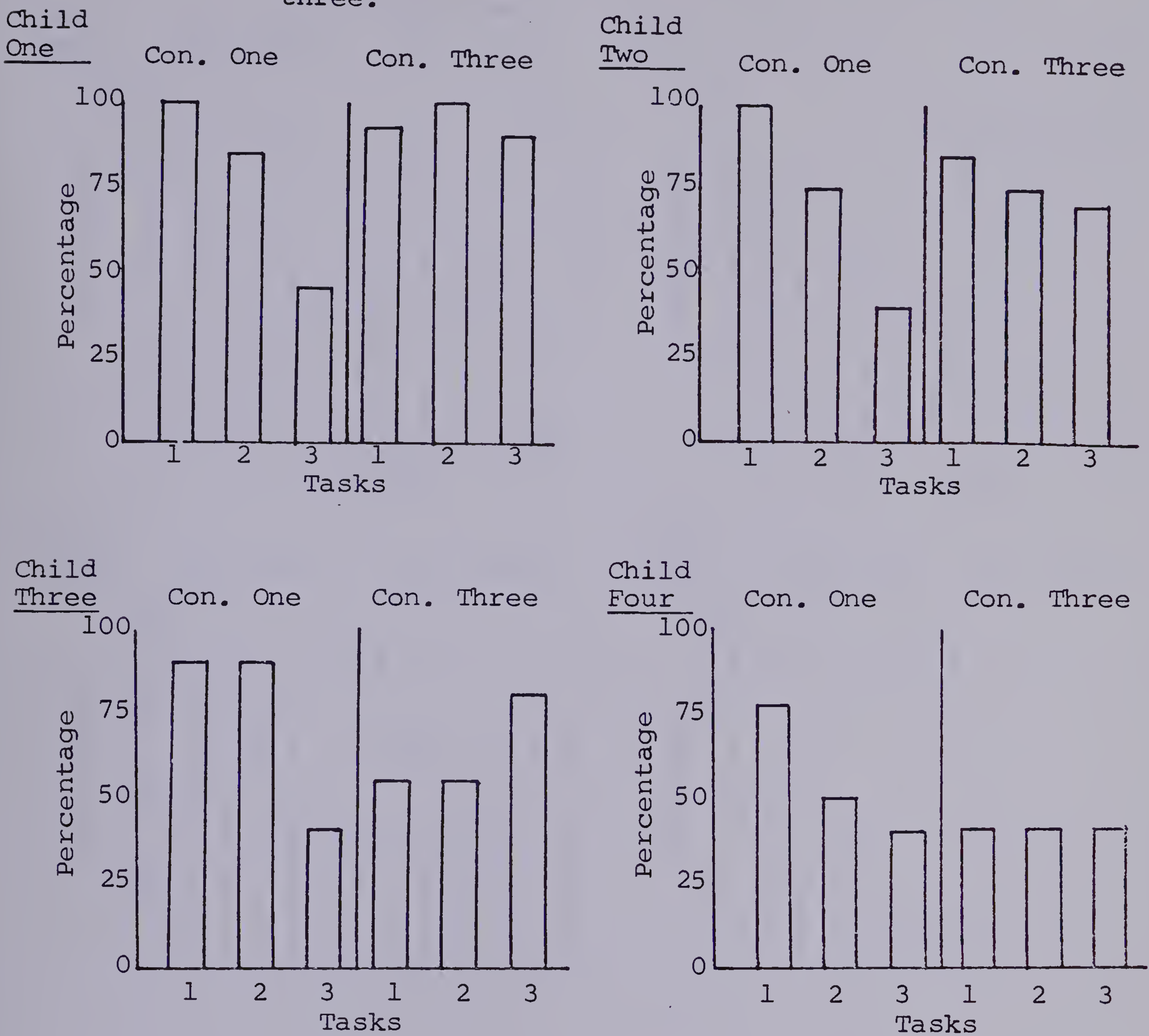
Child four shows greater generalization to GT1 when comparing conditions one with three. This replicates the effect observed in all other children. The remaining generalization results for this subject are at or below the chance level and show no difference between conditions.

Figure twelve shows a comparison of conditions two with four at each generalization task for each child and as an average comparison for all children. There is a replication across all children showing a greater generalization for GT3 in condition four over condition two.

Day One and Day Two

Results are presented for individual children showing

Figure 11: Percentage of generalization to tasks one, two, three. Comparison of condition one to condition three.



Average Percentage of Generalization to Levels One, Two, Three for Conditions One and Three

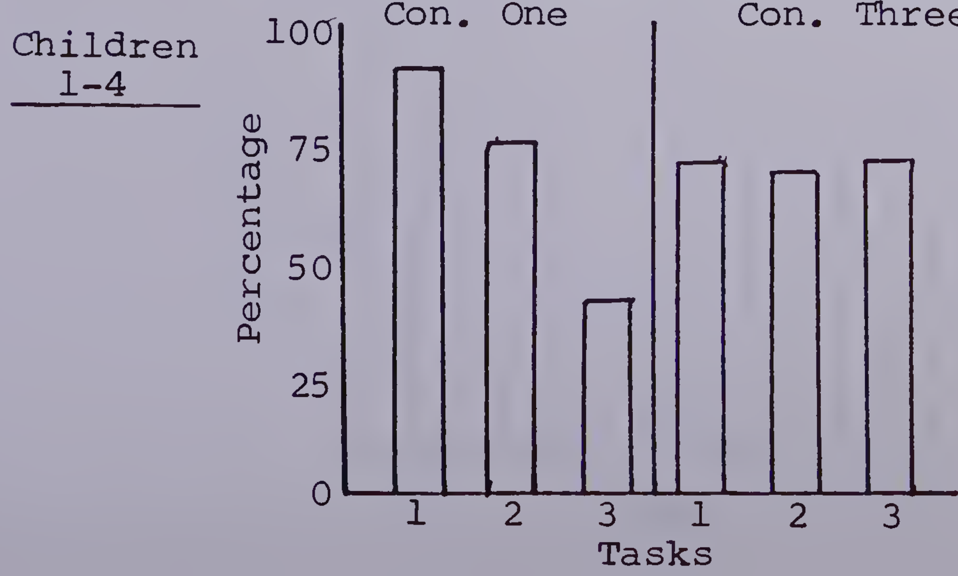
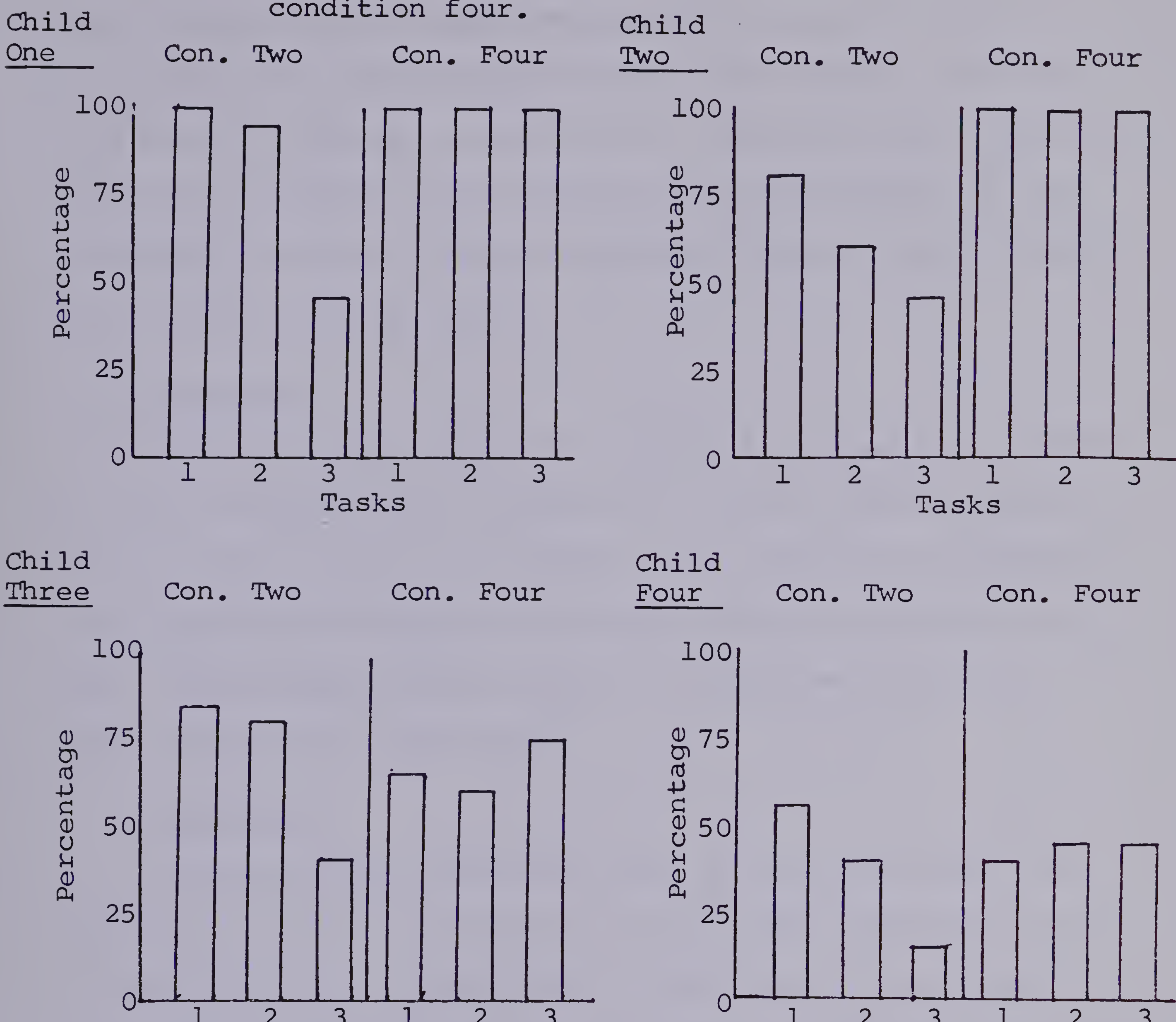
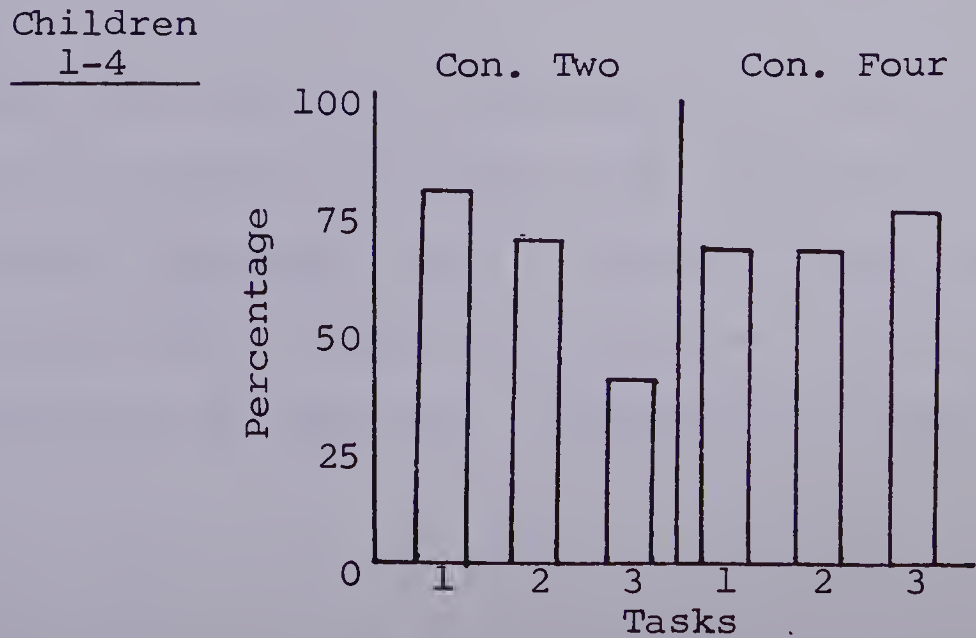


Figure 12: Percentage of generalization to taks one, two, and three. Comparison of condition two with condition four.



Average Percentage of Generalization to Tasks One, Two, Three for Conditions Two and Four



their day one and day two generalization scores to each generalization task collapsed across both members of a concept pair in addition to results for individual pair members. This format allows a closer look at the consistency of generalization results within and between generalization tasks over the two testing days.

Child One:

Figure thirteen shows the results for child one in each of the four experimental conditions. These results show a highly stable pattern of performance to each generalization task across both members of the pair and for individual members. This highly stable pattern is observed within all four experimental conditions.

Child Two:

Figure fourteen shows the results for child two. In condition one there is evidence of a highly stable pattern of results to each generalization task over the two days. Individual concepts do not show a similar trend of stability. "Long" and "short" show a similar trend over days one and two for tasks one and two while task three shows a high degree of variability.

Condition two shows a high level of stability for generalization task two and three with instability at task one. The concept "straight" shows a similar stable trend across days one and two at tasks two and three, with a high level of variability at task one. "Curved" has a stable pattern

Figure 13: Child One.

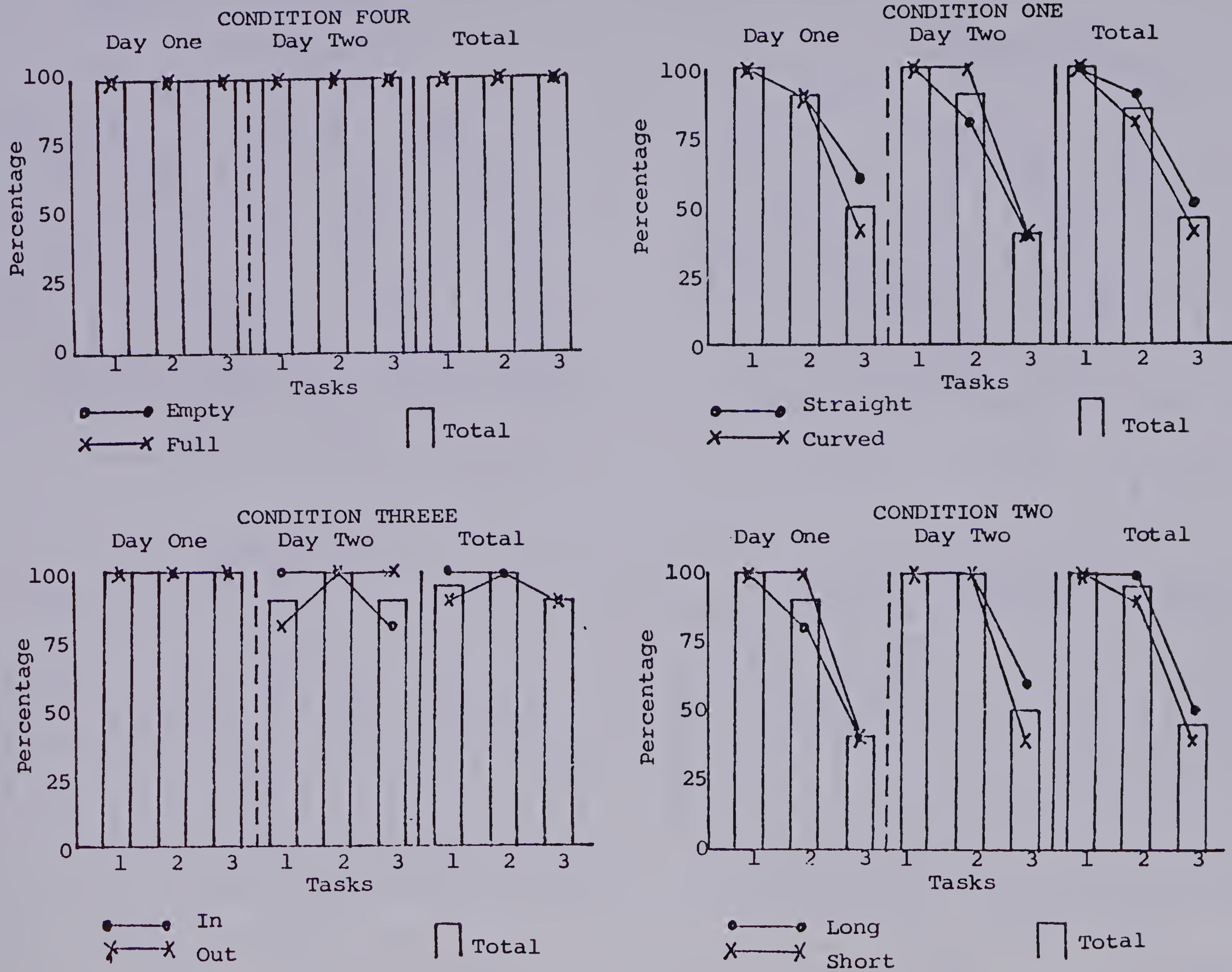
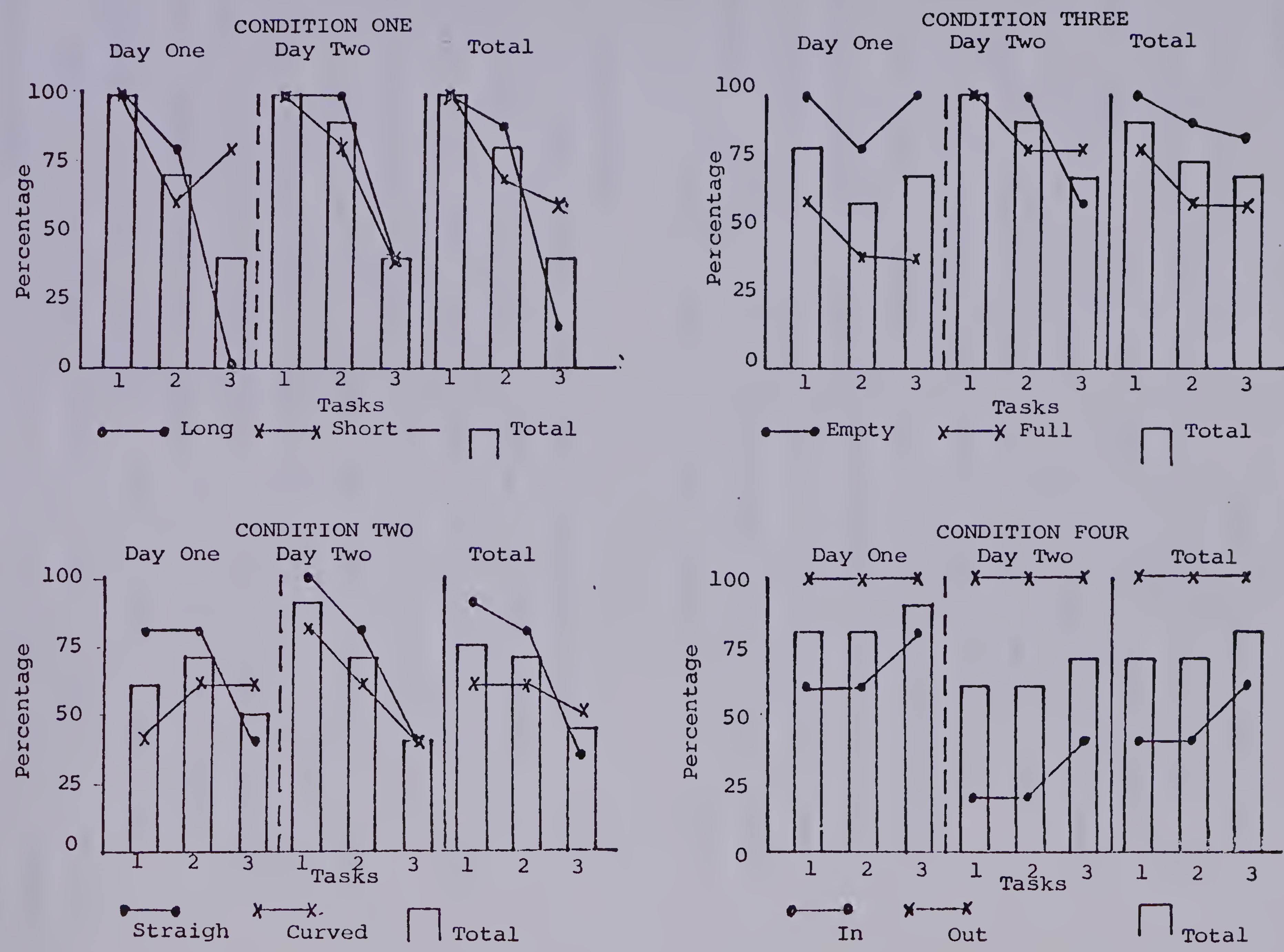


Figure 14: Child Two



over days at task two and a variable pattern for tasks one and three.

Condition three results in a stable pattern for task three across days with variation at tasks one and two. The concept "empty" is stable at task one while instability shows at tasks two and three. "Full" shows a similar trend for tasks one, two and three across days, however, the scores on each task are considerably less in day one.

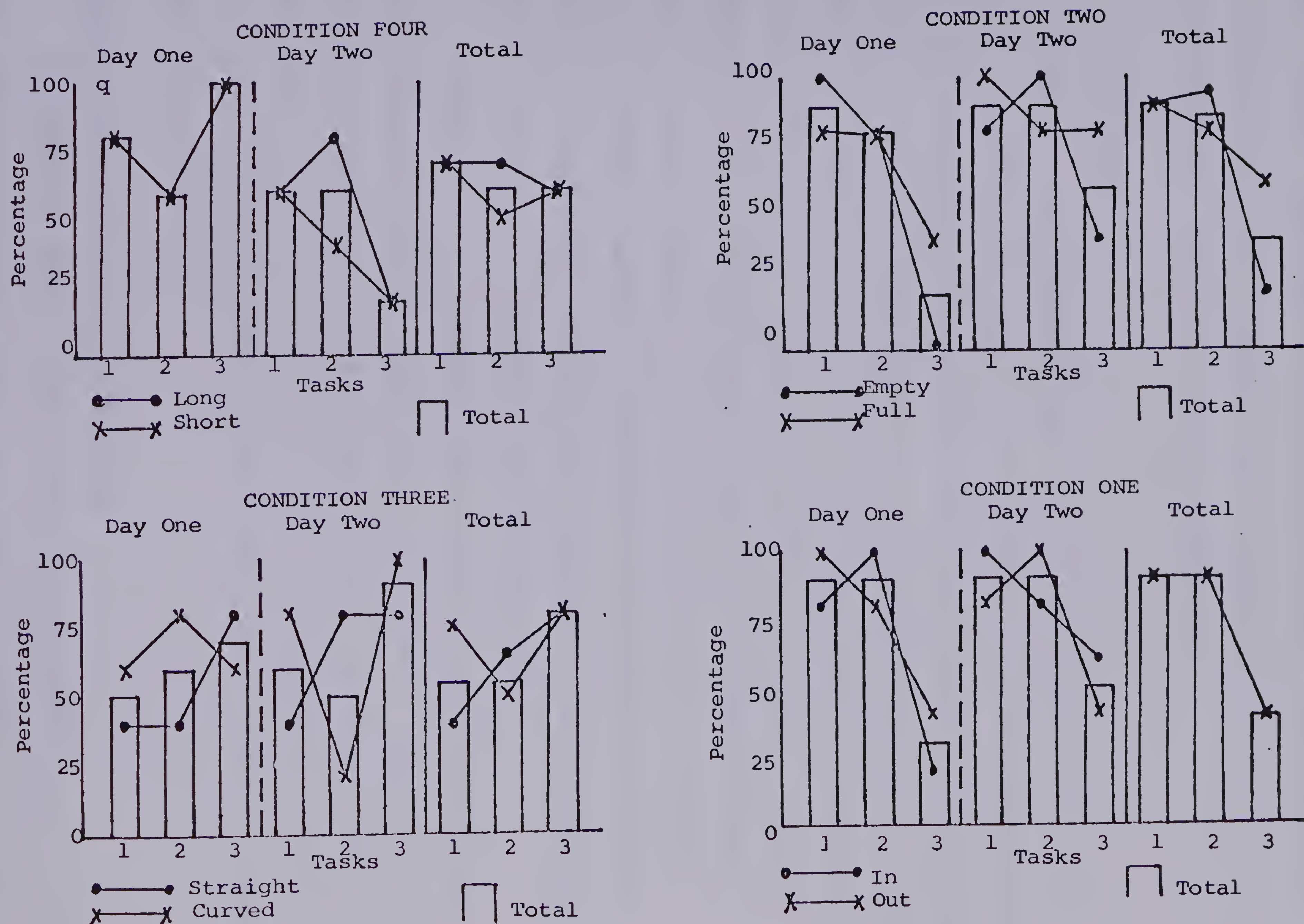
Condition four shows a similar trend across days for individual tasks and each concept member. The scores on each generalization task are lower on day two due to the depressed scores for the concept member "in". The concept "out" shows a highly stable pattern across days one and two.

Child Three:

Individual results for child three are found in Figure fifteen. Condition four shows a stable pattern for generalization task two over days one and two, while tasks one and three in addition to individual concept members show a highly variable pattern over days. Condition three reveals a fairly stable pattern for tasks one and two generalization and not to task three. Concept member "curved" shows a highly unstable pattern across the days, while "straight" is only unstable at generalization task two.

Condition two shows a fairly stable pattern to generalization tasks one and two collapsed over both concept members between days one and two, while task three is unstable. The

Figure 15: Child Three



concept "full" shows stability for generalization task two over the two days while performance on tasks one and three is highly unstable. "Empty" has unstable performance on all generalization tasks over days one and two.

Child Four:

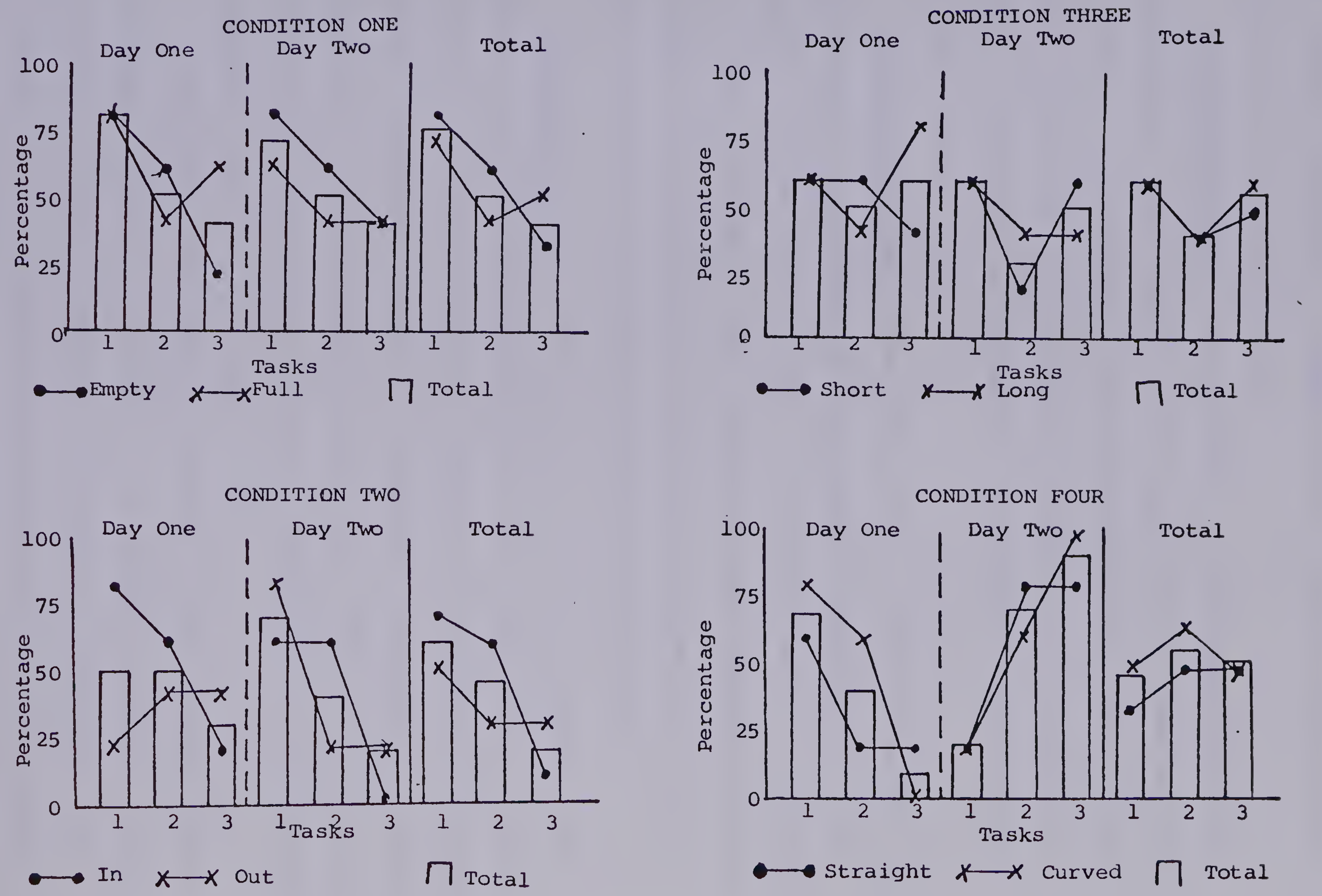
Figure sixteen shows the results for child four. Condition one shows a highly stable pattern of generalization on all generalization tasks across days one and two. Individual concept "empty" shows stability for tasks one and two, while "full" is stable at task two. They show unstable patterns on all other tasks. Condition two shows a highly unstable pattern on each generalization task for collapsed scores across concept members over days one and two. A similar highly unstable pattern is observed to all generalization tasks for individual concept pair members. The only exception is with "in" at generalization task two.

Condition three shows stable results at tasks one and three generalization, while two is unstable. "Long" appears stable in tasks one and two while "short" is stable only at task one. Condition four shows highly unstable results to all generalization tasks and for both concept members across days one and two.

Summary of Generalization Results

Hypothesis Three: The results of group data do not support this hypothesis. There was not an increase in total generalization responses as the number of exemplars increased

Figure 16: Child Four



for either the overall comparison or separate comparisons at each level of IER variation.

Hypothesis Four: Group data and individual subject replication do not support this hypothesis. There was not an increase in total generalization responses to GT1, GT2 or GT3, as the number of exemplars increased for either the overall comparison or separate comparisons at each level of IER variation.

Hypothesis Five: The group results do not support the experimental hypothesis on the overall comparison or the separate comparison at level one NE while weakly supporting it at level two NE. Total generalization responses do not increase as interstimulus variation increases from one to three irrelevant dimensions between concept pairs.

Hypothesis Six: Total generalization responses for generalization tasks one (GT1) and two (GT2) do not show a difference in the results to support the experimental hypothesis at either the overall or separate comparisons. However, responses to generalization task three (GT3) indicates group results which strongly support the hypothesis on the overall comparison and to each level of NE.

Increasing interstimulus variation from one to three irrelevant dimensions results in a) greater generalization responses to GT1 and GT3 for the overall comparison; and b) greater generalization responses to GT3 at NE one and NE two. Consistency of generalization responses observed

over the two testing periods revealed considerable variation for three of the four children. There did not appear to be a general pattern for individual concepts as in some conditions one member of a concept pair was more variable while the opposite member was not. Generally, however, there did appear to be a pattern of more variable performance for concept pairs and generalization task levels over the two testing periods,when trained in experimental conditions three and four.

The relation between acquisition and generalization results are discussed in the next chapter.

CHAPTER VI

DISCUSSION

There are some general statements which can be made from the results of the present study: 1) with the exception of one child in experimental condition four, all children learned all the relational concepts; 2) differential rates for acquisition were found between experimental conditions; 3) differential amounts of generalization were found following acquisition in different experimental conditions; 4) increasing the number of exemplars as one instructional parameter resulted in greater total trials to criterion as the number of exemplars increased from two to four; 5) increasing diversity of exemplars also resulted in greater total trials to criterion as interstimulus variation increased from one to three irrelevant dimensions; 6) number of exemplars as one instructional parameter did not result in increased generalization to generalization tasks one, two or three as number of exemplars increased from two to four; 7) diversity of exemplars as one instructional parameter resulted in an increase in generalization responses to generalization task one and three as the interstimulus variation increased from one to three irrelevant dimensions.

An initial conclusion from these results suggests that diversity rather than number of exemplars is the more crucial

dimension for programming generalization. Condition three (in which there were two exemplars with three irrelevant dimensions varied) produced the greatest generalization for the least amount of instructional time. However, condition three required twice as many total trials to criterion as condition one (in which there were two exemplars with one irrelevant dimension varied. Results from the pilot study showed one child to have only slightly greater total trials to criterion for condition three over condition one, with greater generalization resulting from condition three; while the second pilot child showed fewer total trials to criterion for condition three in addition to greater generalization. During the main experiment this effect was not observed for acquisition, as all children required greater total trials to criterion for condition three over conditions one and two. Generalization results showing a greater generalization from condition three were observed for all experimental children in a similar manner to the pilot children.

The results of the present experiment are discussed in relation to previously reviewed literature, first for acquisition, secondly for generalization and thirdly, for relation between acquisition and generalization.

Acquisition

Previous investigators have found that as stimulus complexity in concept learning tasks is increased on both relevant and irrelevant dimensions, subject performance decreases

(Bourne and Haygood 1959, 1961); while increased stimulus complexity on either relevant or irrelevant dimensions also results in decreased performance (Battig and Bourne 1961). Alternatively, Haygood et al (1970) and Chumbley et al (1971) have shown that increasing stimulus complexity for irrelevant dimensions while relevant dimensions remain constant results in increased performance or fewer performance trials to solution.

The research on increased stimulus complexity for irrelevant dimensions is in conflict when comparing the work of Battig and Bourne (1961) with Haygood et al (1970) and Chumbley et al (1971). Chumbley et al (1971) suggested that differences between their results and those of Battig and Bourne (1961) may be due to the manner in which task complexity was increased in addition to employing different instructional formats.

It is difficult to draw general conclusions concerning the role of irrelevant dimensions and how they influence acquisition as many investigators have employed different concept classes, instructional formats and degrees of increasing task complexity. One possible prediction is that increasing stimulus complexity in a concept learning task, where irrelevant dimensions change and relevant dimensions remain constant may increase the saliency of the relevant dimension and thus facilitate discrimination of the relevant dimension.

The present experiment investigated the influence of

increased stimulus complexity on irrelevant dimensions alone for polar opposite concepts presented in a simultaneous format. As a result, a discussion of the results can only apply to concept class with a similar topography and instructional format.

Stimulus complexity during instructional sets was varied on two different program dimensions. The first was number of exemplars and the second was diversity of exemplars. Logically one could conclude that increasing the number of exemplars from a concept class during instructional sets would result in more trials to criterion. Similarly if the number of exemplars did not increase but the number of irrelevant dimensions for exemplars increased, greater trials to criterion should result. What is not clear from previous research is which of these two program dimensions results in the greater stimulus complexity.

The present study used experimental conditions which increased stimulus complexity in a presumed linear fashion from conditions one through to four. Condition one contained teaching sets with two pairs of exemplars where the irrelevant dimensions (color and form) remained constant. Condition two contained teaching sets with four pairs of exemplars where the irrelevant dimensions form and color remained constant. Condition three contained teaching sets with two pairs of exemplars where the irrelevant dimensions form, color and position varied while condition four contained four pairs of exemplars where the irrelevant dimensions color, form and position varied.

This linear arrangement suggests that number of exemplars as one program dimension will result in less complex teaching sets than diversity of exemplars. The average total trials to criterion across the four experimental children bears this out. Condition one took the least number of average trials to criterion, condition two second, condition three third, and condition four the most. Diversity of exemplars rather than number within instructional teaching sets resulted in greater stimulus complexity as reflected in average number of trials to criterion. Within condition four, diversity and number of exemplars interacted to produce the most complex instructional teaching set.

Stimulus complexity was also observed to increase within each of these two program dimensions. There was an increase in average total trials to criterion as stimulus complexity increased for number of exemplars alone. This was observed at both levels of interstimulus variation as condition one (two exemplars, irrelevant constant) had fewest trials to criterion than condition two (four exemplars, irrelevant constant) and condition three (two exemplars, irrelevant varied) showed less trials to criterion than condition four (four exemplars, irrelevant varied). Similarly there was an increase in average total trials to criterion as stimulus complexity was increased for diversity of exemplars. This was also observed at both levels of number of exemplars as condition one (two exemplars, irrelevant con-

stant) took less trials to criterion when compared to condition three (two exemplars, irrelevant varied) while condition two (four exemplars, irrelevant constant) showed fewer trials to criterion than condition four (four exemplars, irrelevant varied).

These results are viewed as supportive of the research by Bourne and Haygood (1959, 1961) and Battig and Bourne (1961) which suggested that increased stimulus complexity would result in performance decrements. This effect was observed for both program dimensions, number and diversity of exemplars. Diversity as one program dimension resulted in more complex stimulus sets than number, however, the two interacted at high levels of each to produce the most complex stimulus set.

Increasing stimulus complexity on irrelevant dimensions while the relevant dimension remained constant resulted in performance decrements. This suggests that increased stimulus complexity on both program dimensions but primarily diversity of exemplars resulted in instructional sets containing competing irrelevant stimulus cues to which a child may attend. The role of increased stimulus complexity results in greater instructional costs during acquisition, however, the relation to generalization following acquisition has not been investigated. Bourne and Haygood (1959, 1961) and Battig and Bourne (1961) did not investigate the role of increased stimulus complexity during acquisition to generalization.

Generalization

The generalization results from the present study suggest that as task complexity is increased there is also an increase in generalization. However, diversity of exemplars rather than number of exemplars during original learning conditions was the determining factor for increased generalization.

As stimulus complexity was increased through number of exemplars, condition two compared with condition four, there was not any noticeable increase in generalization responses. When stimulus complexity was increased through diversity of exemplars, one irrelevant compared with three irrelevant dimensions, there was a noticeable increase in generalization with a higher level of diversity.

Stokes and Baer (1977) suggested "the optimal combination of sufficient exemplars and sufficient diversity to yield the most valuable generalization is critically in need of analysis" p. 357. These results suggest that the optimal combination for the concept class employed was two exemplars where there were three irrelevant dimensions varied between exemplars presented simultaneously.

Generalization responses were measured in three generalization tasks. Each generalization task had a different degree of stimulus complexity represented by the number of irrelevant dimensions varied between stimulus items. Generalization task one contained items where only position as one

irrelevant dimension was varied, generalization task two contained items where two irrelevant dimensions, position and color, were varied, while generalization task three contained items where three irrelevant dimensions, position, color and form were varied.

Experimental conditions one and two contained stimulus items with interstimulus variation level one (one irrelevant dimension). Generalization test results following conditions one and two showed generalization responses to exemplars at generalization task one and two but not to task three.

Children generalized then to exemplars with a similar level of interstimulus variation encountered during training in addition to exemplars containing an additional level of interstimulus variation (color) not observed during training. However they did not generalize to exemplars containing a combination of two levels of interstimulus variation which was not observed during training (color and form).

Conditions three and four contained stimulus items with interstimulus variation level three (three irrelevant dimensions, color, position, form). Generalization test results following these conditions showed children generalizing to exemplars on each generalization task. Teaching in conditions one and two resulted in generalization to exemplars with one irrelevant dimension varied (task one) or combinations of two irrelevant dimensions varied (task two), but not to items with a combination of three variable irrelevant dimensions (task three). Teaching in conditions

three and four resulted in generalization exemplars with one (task one), two (task two) and three (task three) variable irrelevant dimensions.

Modigliani and Rizza (1971) suggested that there is no reason to assume that subjects will generalize to items which do not contain the same number of relevant and irrelevant dimensions observed during training. Modigliani (1971), Modigliani and Rizza (1971) found that in the presence of the cue defining a concept, generalization is governed by the nature of both relevant and irrelevant information. Modigliani and Rizza (1971) stated "the presence of a defining attribute is a necessary but not sufficient condition for the identification of a stimulus as an instance of the concept" p. 239. In a similar vein Stokes and Baer (1977) suggested that generalization may be a function of the stimulus characteristics during training.

The results from the present study would support the statements by Modigliani and Rizza (1971) and Stokes and Baer (1977). Children did not generalize, with the exception of color, to novel items which contained the relevant cue they were trained on and irrelevant dimensions they were not exposed to. The irrelevant dimensions they were not exposed to acted as competing stimuli and did not facilitate generalization. However, if the child was exposed to the irrelevant dimension, he was able to identify novel exemplars in the presence of those irrelevant dimensions.

Generalization to novel exemplars was only observed

if the child had prior exposure to the relevant dimension in association with the irrelevant dimensions contained in the novel exemplars. The exception to this was responding to generalization task two items following training conditions one and two. Here the child was not exposed to the irrelevant dimension color in association with the relevant dimension. This may be explained by the degree of irrelevant information changing between exemplar pairs within a teaching set.

However, form also changed between stimulus pair items as all pairs in conditions one and two were different. This did not assist the children in generalization to generalization task three.

An alternative explanation is that color was not attended to as part of the solution, however, form was. Then when form was varied within stimulus pairs during generalization task three, the child became confused, as previously form was constant. This change in the arrangement of form as an irrelevant dimension produced a learning context to which the child had not previously been exposed. As a result the child did not know the correct solution given the new arrangement of irrelevant cue dimensions. Generalization to novel items then, occurred primarily as a function of original training conditions.

Relation between Acquisition and Generalization

Stokes and Baer (1977) suggested two research questions:

1) is the best procedure to train many exemplars with little diversity at the outset and then expand the diversity to include dimensions of the desired generalization?

2) is it a more productive endeavor to train fewer exemplars that represent a greater diversity and persist in the training until generalization emerges? p. 357.

As the present investigation did not employ a cumulative training model it is difficult to make any direct statement from the results concerning the two points raised by Stokes and Baer (1977). However, the results do bring up some interesting points concerning these two research issues.

It was shown that diversity rather than number of exemplars was the crucial dimension for facilitating generalization. Condition three, containing two exemplars with three irrelevant dimensions produced the greatest generalization with the least amount of instructional time. However, condition three required twice as many average total trials to criterion as condition one which contained two exemplars with one irrelevant dimension varied.

This increase in total trials to criterion for condition three resulted in approximately seventy percent more generalization for generalization task three. The question to be asked is: "what would have happened if cumulative teaching was introduced where the concept pairs were taught to criterion in condition one, followed by training on condition three?" Thus, cumulative introduction of irrelevant dimensions may have resulted in fewer trials to criterion than

condition three alone, while also showing similar generalization.

It appears from the results that the children generalized to novel items which contained similar conditions during original learning. As a result a cumulative strategy would appear to be an appropriate approach. The present results unfortunately only suggest an optimal combination for the dimensions number and diversity of exemplars which were defined by the experimental conditions. Although the results point to diversity as the crucial dimension for programming for generalization there are many unanswered questions in the form of different ways that diversity of exemplars can be programmed.

Suggestions for Future Research

There are two major implications from the present research which need further investigation. The first concerns the role of diversity of exemplars as an instructional dimension when programming for generalization. In the present experiment it was not clear if the greater generalization to task three for condition three over condition one was due to increased interstimulus complexity or simply due to more instructional time. This question could be investigated by comparing the total trials to criterion and resultant generalization for concept pairs taught in condition three alone versus concept pairs taught to criterion in condition one followed by condition three. The combination of conditions one and three would increase stimulus complexity in a serial

manner rather than training continuously at a higher level of complexity. Another variation would be to include a teaching set which contains exemplars from both conditions one and three rather than presenting training items in a serial manner for condition one followed by condition three. A third variation would be to include the same exemplars within a teaching set but interchanging the common irrelevant dimensions. For example, train a "long-short" pair for pencils and dogs which are a different color (condition one); then following criterion, retain the same examples but interchange the irrelevant dimension form. The exemplars now become a long dog-short pencil, long pencil-short dog, with position also varied within the instructional set.

Secondly, the present research employed a specific class of polar opposite concepts presented in a simultaneous format. As a result the generalization of the present findings may not be appropriate for other concept classes and presentation formats. Stokes and Baer (1977) suggested that program dimensions for generalization may be different or combine differently for separate concept classes. It would be important to investigate the role of diversity of exemplars when programming for generalization with other concept classes and presentation formats.

It is also suggested that the present study and results are in need of further replication with additional children both mentally handicapped and of normal intelligence. Such a replication would serve to enhance or limit generalization

of the results. In addition, such further research may serve to highlight important differences between these two populations when designing programs for generalization.

Some Implications for Instructional Programming

The instructional procedure employed in the present research trained both members of a concept pair together. During any teaching set a child had equal exposure to and instruction with both members of the concept pair.

Two interesting behavior patterns which were observed during training may suggest that this is not the optimal way to train. The first deals with the influence of overlearning observed for two children. In this situation one member of a concept pair (i.e. long) reached criterion prior to the opposite member (i.e. short). As a result, due to the instructional procedure one concept pair member received additional learning following an individual criterion being met, until the opposite member also reached criterion.

It was observed that once this situation of overlearning arose the opposite concept member very quickly also reached criterion. In addition, the concept pair member experiencing overlearning also showed slightly greater generalization than its opposite member. An alternative instructional procedure than would be to train one member of a concept pair to criterion then introduce the second member to instruction. This may result in quicker acquisition for the first concept taught while also facilitating acquisition for the second

concept. Given a situation where the same teaching sets are used for both concept members then initial training for one member should facilitate discrimination of the second. This suggests an instructional procedure where relational concepts are taught in a serial rather than concurrent manner.

Panyan and Hall (1978) found no difference between serial and concurrent training procedures during acquisition, however, there was greater generalization following concurrent training. An interesting research question then, would be to investigate the effect of serial and concurrent training procedures for relational concepts during acquisition and resultant generalization.

The second observation is concerned with spontaneous language usage during training. The instructional procedure used teaching sets with a prearranged order of introduction and request for individual concepts. It was noticed that three of the four children began to spontaneously label and point to a stimulus card before they were requested to point to either card.

Given that labelling of the correct concept was not required it was ignored. However, more important is the situation where the trainer is going to ask the child to touch one concept member (i.e. long) and the child spontaneously before any instruction points to the opposite concept member (i.e. short) and says "short". Here the child's correct spontaneous verbal and nonverbal behavior is ignored while

the trainer requests a nonverbal response for the opposite concept.

Given that these three children did not have these labels in their verbal behavior prior to the experiment, it seemed unfortunate not to reinforce it in some way. The children most probably acquired the labels through hearing the instructor use them during instruction. Here is a situation of language learning which was not knowingly programmed. Unfortunately, the instructional procedure did not allow for these spontaneous verbalizations to be reinforced.

A serial rather than concurrent training procedure would allow a trainer to reinforce spontaneous correct labelling. In a serial training sequence one concept is trained alone before the other. Therefore a trainer could reinforce labelling if it occurred during initial training of one concept pair member and during training for the second. This would allow reinforcement of expressive language usage when the instructional procedure is initially for receptive language.

References

Anderson, D. R., Godson, G. D., and Willand, J. G. Instructional programming for the handicapped student. Springfield, Illinois: Charles C. Thomas, 1976.

Anderson, Richard C., and Faust, Gerald W. Educational psychology: the science of instruction and learning. New York: Dodd, Mead and Col, 1974.

Archer, E. J. Concept identification as a function of obvious relevant and irrelevant information. Journal of Experimental Psychology, 1962, 63, 616-620.

Baine, David. Some problems associated with achievement testing of the mentally retarded. Mental Retardation Bulletin, 1977, 5, 49-61.

Battig, W. F., and Bourne, L. E., Jr. Concept identification as a function of intra- and inter-dimension variability. Journal of Experimental Psychology, 1961, 61, 329-333.

Becker, W., Engelmann, S., and Thomas, D. Teaching: A course in applied psychology. Chicago: Science Research Associates, 1971.

Becker, W. C., Engelmann, S., and Thomas, D. R. Teachings: Cognitive learning and instruction. Toronto: Science Research Associates, 1975.

Becker, W. C., and Englemann, S. Teaching 3: Evaluation of instruction. Toronto: Science Research Associates, 1977.

Bellamy, G. T., and Bellamy, T. T. Descriptive concepts for preschool retarded children. Education and Training of the Mentally Retarded, 1974, 9, 115-121.

Bijou, S. W., and Baer, D. M. Child development, Vol. 1. A systematic and empirical theory. New York: Appleton-Century-Crofts, 1961.

Bolfon, Neil. The psychology of thinking. London: Methuen and Co. Ltd., 1972.

Bourne, L. E., Jr. Human conceptual behavior. Boston: Allyn and Bacon, 1966.

Bourne, L. E., Jr. Knowing and using concepts. Psychological Review, 1970, 70, 546-556.

Bourne, Lyle E. Jr., and Dominowski, Roger L. Thinking. _Annual Review of Psychology_, 1972, _23_, 207-276.

Bourne, L. E., Jr., and Guy, D. E. Learning conceptual rule, II: The role of positive and negative instances. _Journal of Experimental Psychology_, 1968, _77_, 488-494.

Bourne, L. E., Jr., and Jennings, P. C. The relationship between contiguity and classification learning. _Journal of Genetic Psychology_, 1963, _69_, 335-338.

Bricker, D., and Bricker, W. A. Toddler Research and Intervention Project Report - Year II. _IMRID Behavioral Science Monograph_, John F. Kennedy Center for Research on Education and Human Development, George Peabody College, 1972, No. 21.

Bricker, William, A. Identifying and modifying behavioral deficits. _American Journal of Mental Deficiency_, 1970, _75_, 16-21.

Bruner, J. S., Goodman, J. J., and Austin, G. A. _A study of thinking_. New York: Wiley, 1956.

Chumbley, J., Lau, Portia, Roy, D., and Haile, G. Concept identification as a function of intradimensional variability, availability of previously presented material, and relative frequency of relevant attributes. _Journal of Experimental Psychology_, 1971, _90_, 163-165.

DeCecco, John, P., and Crawford, William R. _The psychology of learning and instruction._ 2nd ed. Englewood Cliffs, New Jersy: Prentice Hall Inc., 1974.

Deese, J., and Hulse, S. H. _The psychology of learning_(3rd ed.) New York: McGraw-Hill, 1967.

Ellis, Henry C. _Fundamentals of human learning and cognition_. Dubuque, Iowa: Wm. C. Brown Co., 1972.

Englemann, S., The effectiveness of direct instruction on IQ performance and achievement in reading and arithmetic. In J. Hellmuck (ed.), _Disadvantaged children_, Vol. 3, New York: Brunner/Mazel, 1970.

Engelmann, S., and Osborn, J. _Distar R Language level I_. Chicago: Sicence R search Associates, 1976.

Fields, P. E. Studies in concept formation: I. The development of the concept of triangularity by the white rat. _Comparative Psychological Monographs_, 1932, _9_, whole #42.

Flavell, John H. Concept development in Paul H. Mussen (Ed.). Carmichael's manual of child psychology vol. one, 3rd edn. New York: John Wiley and Sons, Inc., 1970.

Fredericks, A. D., Riggs, C., Furey, T., Grove, D., Moore, W., McDonnel, J., Jordon, E., Hanson, W., Baldwin, V., and Wadlow, M. The teaching research curriculum for moderately and severely handicapped. Springfield, Illinois: Charles C. Thomas, 1976.

Gagne, Robert M. The learning of principles in Klausmeir, Herbert J., and Harris, Chester W. (Eds.). Analysis of concept learning. New York: Academic Press, 1966.

Gagne, Robert M. The conditions of learning. 3rd Edition. New York: Holt, Rinehart and Winston, 1977.

Goldiamond, I. Perception. In A. J. Bachrach (Ed.). Experimental foundations of clinical psychology. New York: Basic Books, 1962.

Green, E. J. Concept formation: a problem in human operant conditioning. Journal of Experimental Psychology, 1955, 49, 175-180.

Hayden, A. H., and Haring, N. G. The acceleration and maintenance of developmental gains in Down's Syndrome school-age children, In Peter Mittler (Ed.), Research to practice in mental retardation, Vol. I, Care and intervention. Baltimore: University Park Press, 1977.

Haygood, R. C., Harbert, T. L., and Omlor, J. A. Intradimensional variability and concept identification. Journal of Experimental Psychology, 1970, 83, 216-219.

Haygood, R. C., and Stevenson, M. Effects of number of irrelevant dimensions in non-conjunctive concept learning. Journal of Experimental Psychology, 1967, 74, 302-304.

Horton, David L., and Turnage, Thomas W. Human Learning. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1976.

Hovland, C. I. A "communication analysis" of concept learning. Psychological Review, 1952, 59, 461-472.

Hunt, E. B. Concept learning: An information processing problem. New York: Wiley, 1962.

Kazdin, A. E. Statistical analysis for single-case experimental designs. In M. Hasin and D. Barlow, (Eds.), Strategies for studying behavior change. New York: Pergamon Press, 1976.

Keller, F. S., and Schoenfeld, W. N. Principles of psychology. New York: Appleton-Century-Crofts, 1950.

Kendler, H. H. Environmental and cognitive control of behavior. American Psychologist, 1971, 26, 962-973.

Kendler, Tracy S. Concept formation. Annual review of psychology. Palo Alto, California: Annual Reviews Inc., 1961, 447-473.

Kendler, Tracy S., and Kendler, H. H. Reversal and nonreversal shifts in kindergarten children. Journal of Experimental Psychology, 1959, 58, 56-60.

Klausmeier, Herbert J., and Harris, Chester W. (Eds.), Analysis of concept learning. New York: Academic Press, 1966.

Kurtz, K. H., and Hovland, C. I. Concept learning with different sequences of instances. Journal of Experimental Psychology, 1956, 66, 353-366.

Kysela, G. M., Daly, K., Hillyard, A., McDonald, L., Butt, B., Ahisten, J., McDonald, S., and Smith, N. The early education project: I. Paper presented to the Canadian Psychological Association, Toronto, 1976.

Kysela, G. M. A home and school based program for infants and preschool children exhibiting developmental handicaps. Final report submitted to Alberta Education Early Childhood Services, Planning and Research. September 1, 1977 (in press).

Miller, Keith L. Principles of everyday behavior analysis. Monterey, California: Brooks/Cole Publishing Co., 1977.

Mittler, Peter (Ed.) Research to practice in mental retardation Volume I. Care and intervention. Baltimore: University Park Press, 1977.

Modigliani, V. On the conservation of simple concepts: Generality of the affirmation rule. Journal of Experimental Psychology, 1971, 87, 234-240.

Modigliani, V., and Rizza, J. P. Conservation of simple concepts as a function of deletion of irrelevant attributes. Journal of Experimental Psychology, 1971, 90, 280-286.

Namikas, G. Concept identification as a function of relevance of pretraining and percentage of information feedback. Psychonomic Science, 1967, 8, 261-262.

Nelson, K. Concept, word, and sentence: Interrelations in acquisition and development. Psychological Review, 1974, 81, 267-285.

Osgood, C. E. Method and theory in experimental psychology. New York: Oxford University Press, 1953.

Panyan, Marion, C., and Hall, Vance R. Effects of serial versus concurrent task sequencing on acquisition, maintenance, and generalization. Journal of Applied Behavior Analysis, 1978, 11, 67-74.

Price, H. H. Thinking and experience. 2nd ed. London: Hutchinson, 1962.

Risley, T. R. Behavior modification: An experimental-therapeutic endeavor. In L. A. Hammerlynck, P. O. Davidson, and L. E. Acker, (Eds.), Behavior modification and ideal mental health services. Calgary, Canada: University of Clagary Press, 1970, 103-127.

Schultz, Jr., and Dodd, D. H. Intradimensional variability in concept identification: A replication extension and partial clarification of the Haygood, Harbert, and Omlor findings. Journal of Experimental Psychology, 1972, 94, 321-325.

Sidman, M. Tactics of scientific research: Evaluating experimental data in psychology. New York: Basic Books, 1960.

Sidman, M., and Cresson, O. Reading and cross-modal transfer of stimulus equivalents in severe retardation. American Journal of Mental Deficiency, 1973, 77, 515-523.

Skinner, B. F. Science and human behavior. New York: Macmillan, 1953.

Spradlin, J. E., Cotter, V. W., and Braxley, N. Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. American Journal of Mental Deficiency, 1973, 77, 556-566.

Spradlin, J. E., and Dixion, M. H. Establishing conditional discriminations without direct training: Stimulus classes and labels. American Journal of Mental Deficiency, 1976, 80, 555-561.

Stokes, Trevor F., and Baer, Donald M. An implicit technology of generalization. Journal of Applied Behavior Analysis, 1977, 10, 349-368.

Sulzer-Azaroff, Beth, and Mayer, G. Roy. Applying behavior-analysis procedures with children and youth. New York: Holt, Rinehart and Winston, 1977.

Terrace, H. S. Discrimination learning with and without errors. _Journal of Experimental Analysis of Behavior_, 1963, _6_, 21-27.

Tjossem, T. D. (Ed.), _Intervention strategies for high risk infants and young children_. Baltimore: University Park Press, 1976.

Walker, C. M., and Bourne, L. E., Jr. Concept identification as a function of amounts of relevant and irrelevant information. _American Journal of Psychology_, 1961, _74_, 410-417.

APPENDIX A

Pilot Study

Pilot Study

A pilot study was conducted with two children prior to completion of the experiment with four additional children. The results of the pilot study are presented in Figures P1 - through P6.

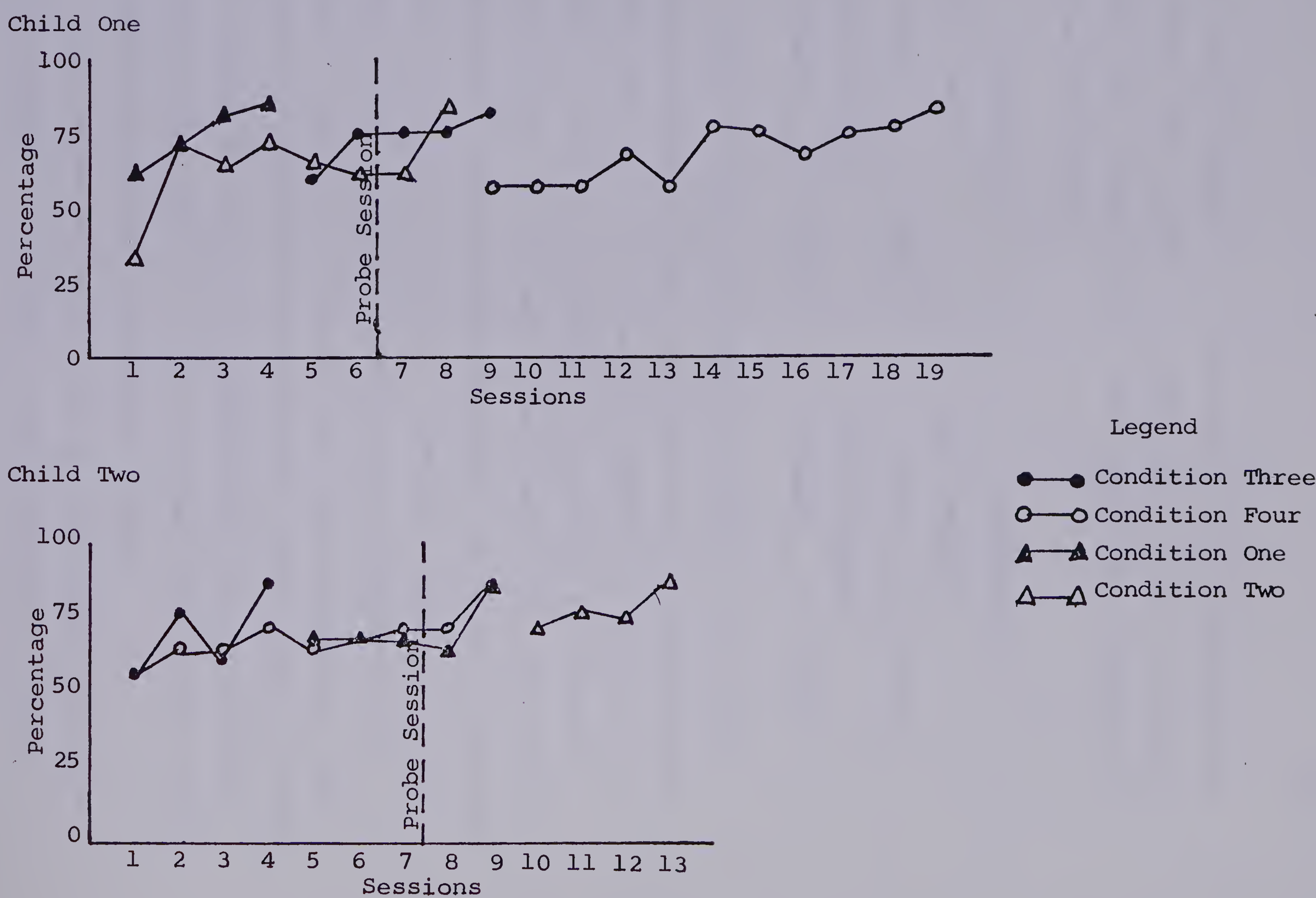
Figure P1 shows the introduction of experimental conditions and percentage correct each session for children one and two. During the pilot study it was observed that the experimental procedure may be teaching a response set. A probe session was initiated for each child to test if a response set was present. Probe sessions occurred between session six and seven for child one and between sessions seven and eight for child two.

Probe Session

The original procedure consisted of presenting both members of the concept pair simultaneously and first requesting a response to one member and, secondly, requesting the opposite member without changing their position or interchanging new stimulus cards. As a result, if a child chose the right position and was correct, the next choice for a correct response was automatically on the left. Each child was observed responding to the second card (position) immediately following feedback on their first choice and prior to any instruction by the experimenter.

A probe session was initiated to both subjects to determine if a response set was being produced. Condition A was

Figure P1: Percentage Correct each Session



the original procedure for the second trial and condition B changed the procedure by repeating a request to respond to the first concept rather than asking for a response to the second concept.

Condition A	Condition B
First Request	First Request
"Touch _____" (long)	"Touch ______" (long)
Second Request	Second Request
"Touch _____" (short)	"Touch _____" (long)

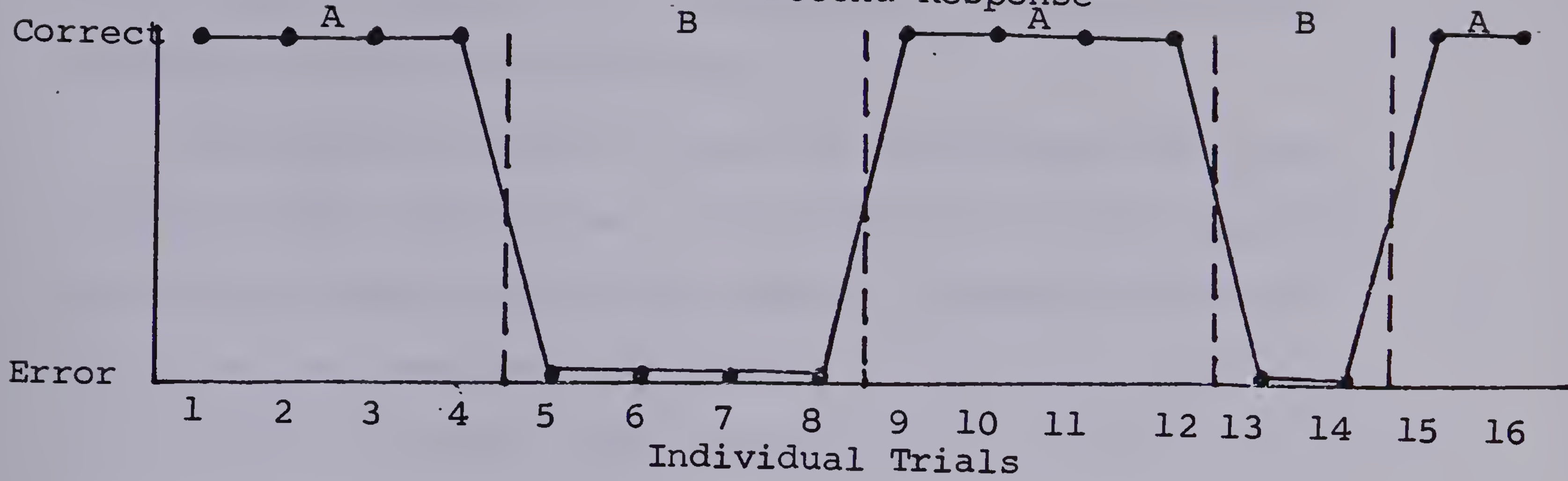
Figure P2 shows the individual data for each child during the probe session. The graphs show individual performance for each child's second response in the two conditions. The sixteen trials for each child are shown as either correct or incorrect. The results show both children responded at 100% for second responses during the A condition and 0% during the B condition. Incorrect responses to the second request in the B condition were followed with the standard correction procedures and retrial procedure. Both children responded correctly following the correction procedure; however, the correction procedure was not strong enough to change their responding for the first instructional trial to the second concept on subsequent trials.

The response set established by the original procedure was too strong to be altered by the correction procedure during one probe session. At this point the original procedure was continued for both pilot children allowing them

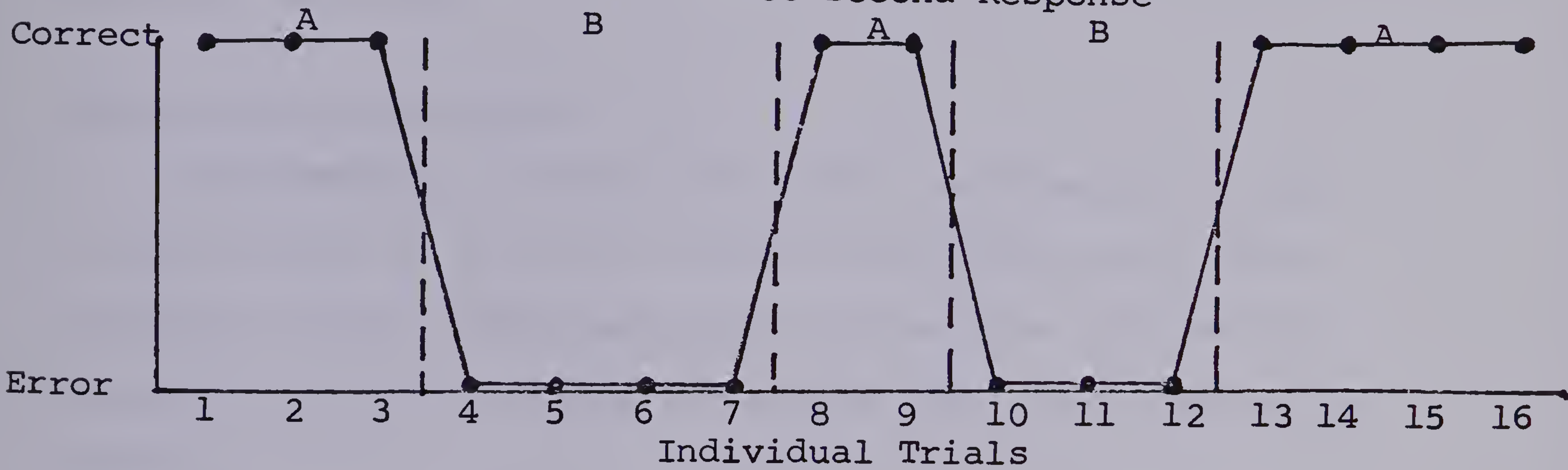
Figure P2: Probe Sessions

Child One

Condition Two Second Response

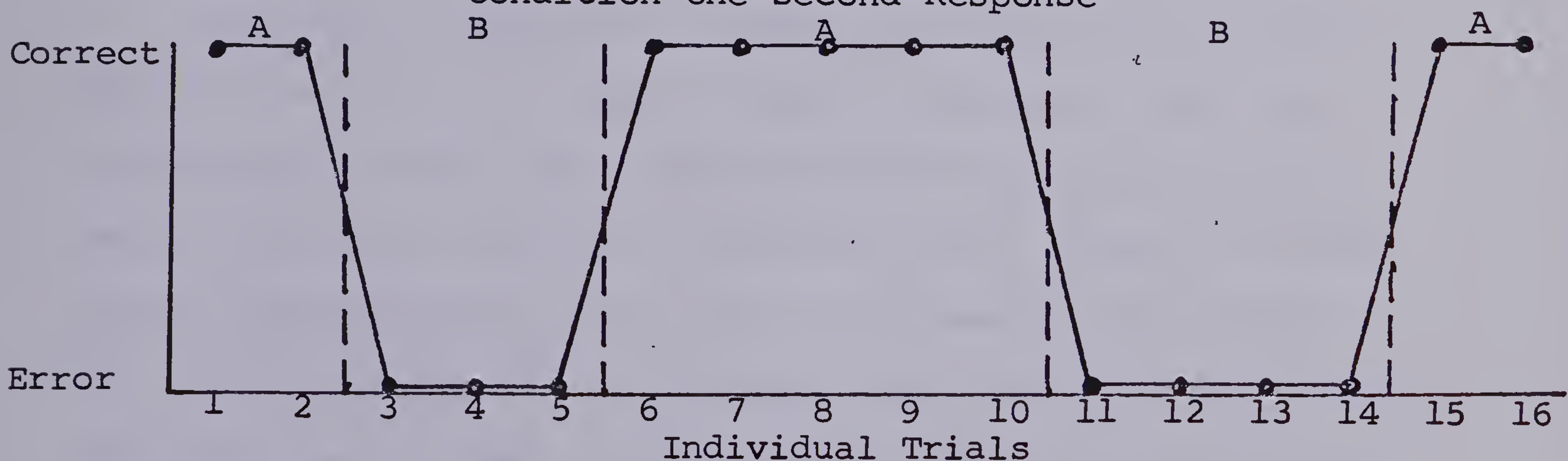


Condition Three Second Response

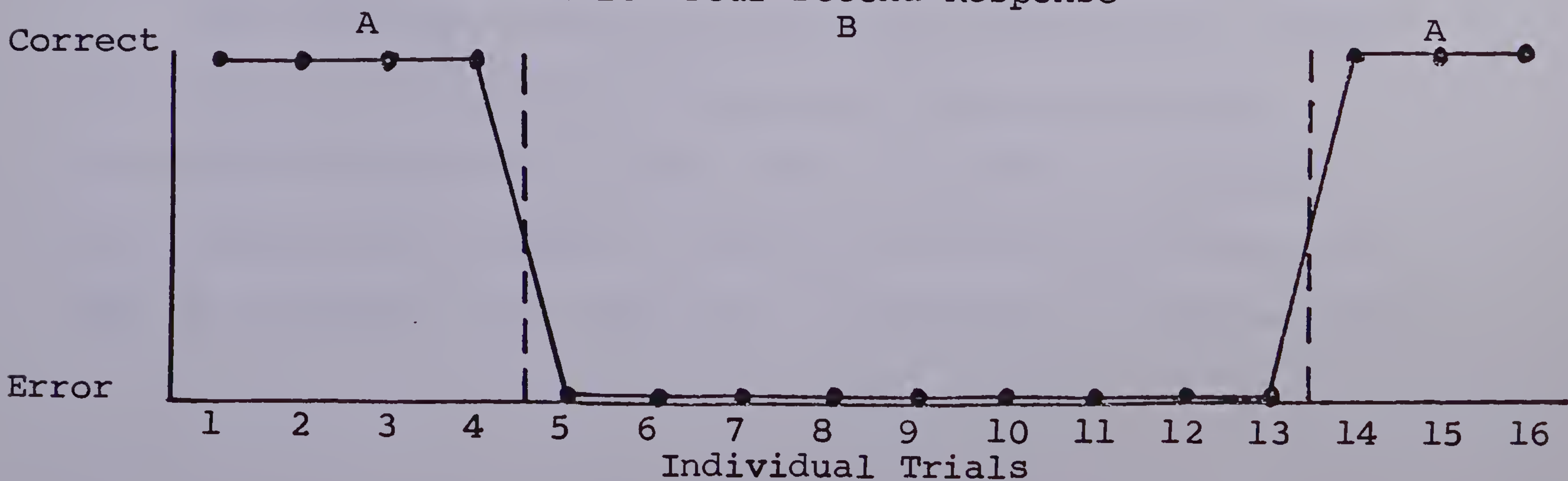


Child Two

Condition One Second Response



Condition Four Second Response



to finish the experiment. The experimental procedure was changed for the main experiment.

No further attempt was made to investigate the topography of the response set. The inability of the correction procedure to gain instruction control suggests that each child may not have been attending to either the instruction or stimulus cards but rather displaying an inappropriate conditioned response.

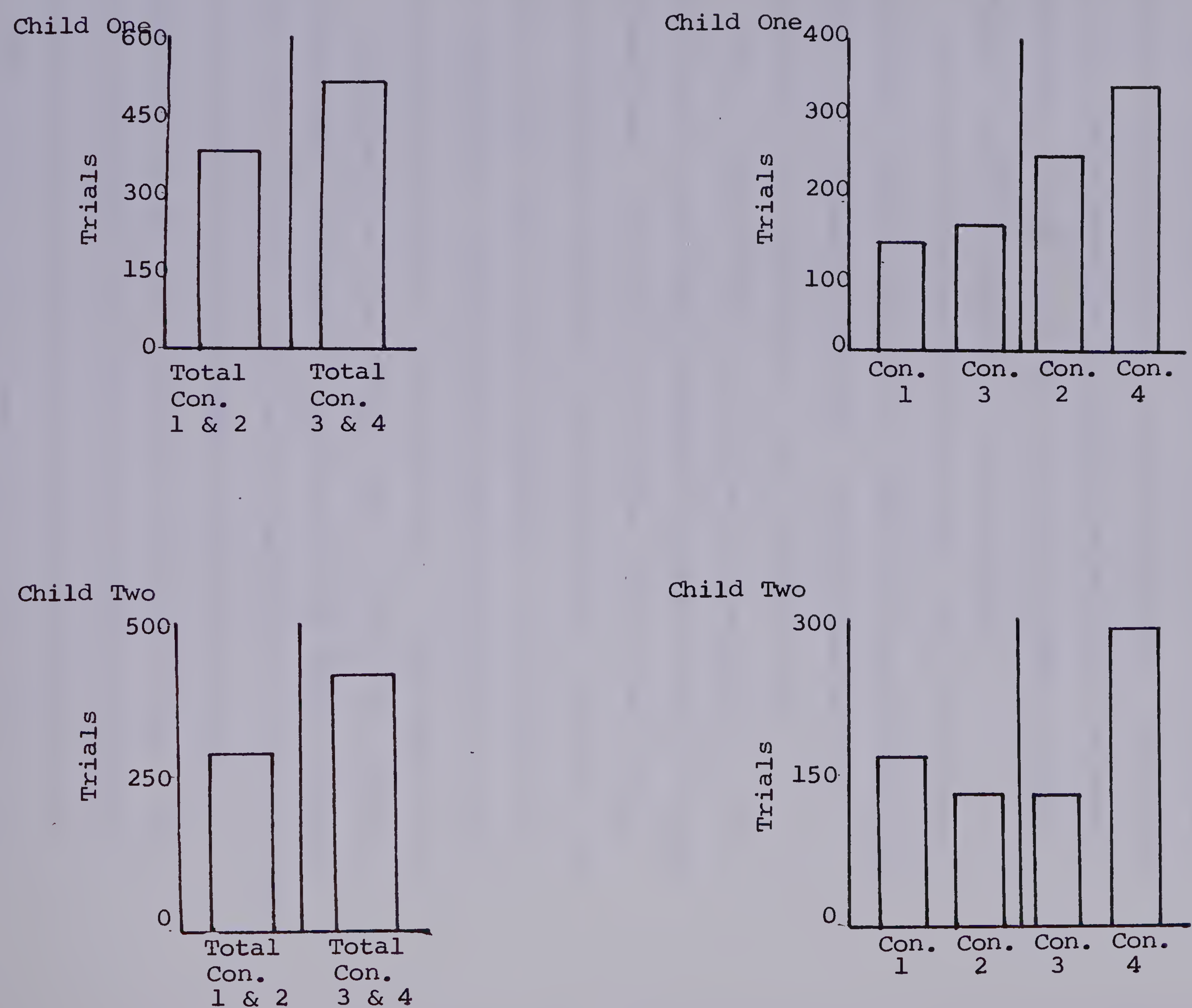
Results and Discussion

Hypothesis one stated that there would be an increase in the number of trials to criterion as the interstimulus variation (IER) increases from level one (one irrelevant dimension varied) to level two (three irrelevant dimensions varied).

An overall comparison between IER levels one and two is made by comparing the total trials to criterion (TTC) for conditions one and two compared with total trials to criterion (TTC) for conditions three and four. Figure P3 shows these comparisons for each child. The comparison of TTC for conditions one and two totals with three and four shows both children to have greater TTC for three and four totals.

Two additional comparisons were made showing the effect of increasing interstimulus variation from one to three irrelevant dimensions at each level of number of exemplars. This comparison is shown in Figure P3 for both children with TTC for condition one (IER level 1, NE Level 1) compared with

Figure P3: Total trials to criterion for conditions one and two total, three and four total; condition one with two and three with four.



condition three (IER level 2, NE level 1) and condition two (IER level 1, NE level 2) compared with condition four (IER level 2, NE level 2).

The results show for child one that increasing inter-stimulus variation from one to three irrelevant dimensions at level one NE results in slightly greater TTC for condition three. Child two on the other hand, shows a slight increase on TTC for condition one over three.

At level two NE both children show that as irrelevant dimensions increase from one to three irrelevant dimensions; there is an increase in the TTC. This is observed in a greater TTC for condition four over condition two.

These results from two children are supportive of the hypothesis on the overall comparison and at level two number of exemplars. That is, as the number of irrelevant dimensions increases from one to three there was a resultant increase in the total trials to criterion. This effect was not observed for increasing irrelevant dimensions at level one NE and was viewed as not supportive of they hypothesis.

Hypothesis two stated that there would be an increase in the number of trials to criterion as the number of exemplars (NE) increased. The overall comparison for this hypothesis is made by taking the total trials to criterion for conditions one and three compared with two and four. This shows the influence of increasing the number of exemplars from two to four at each level of interstimulus variation.

The results of this comparison for both children are

presented in Figure P4. These results clearly show that conditions two and four totals for both children were greater than conditions one and three.

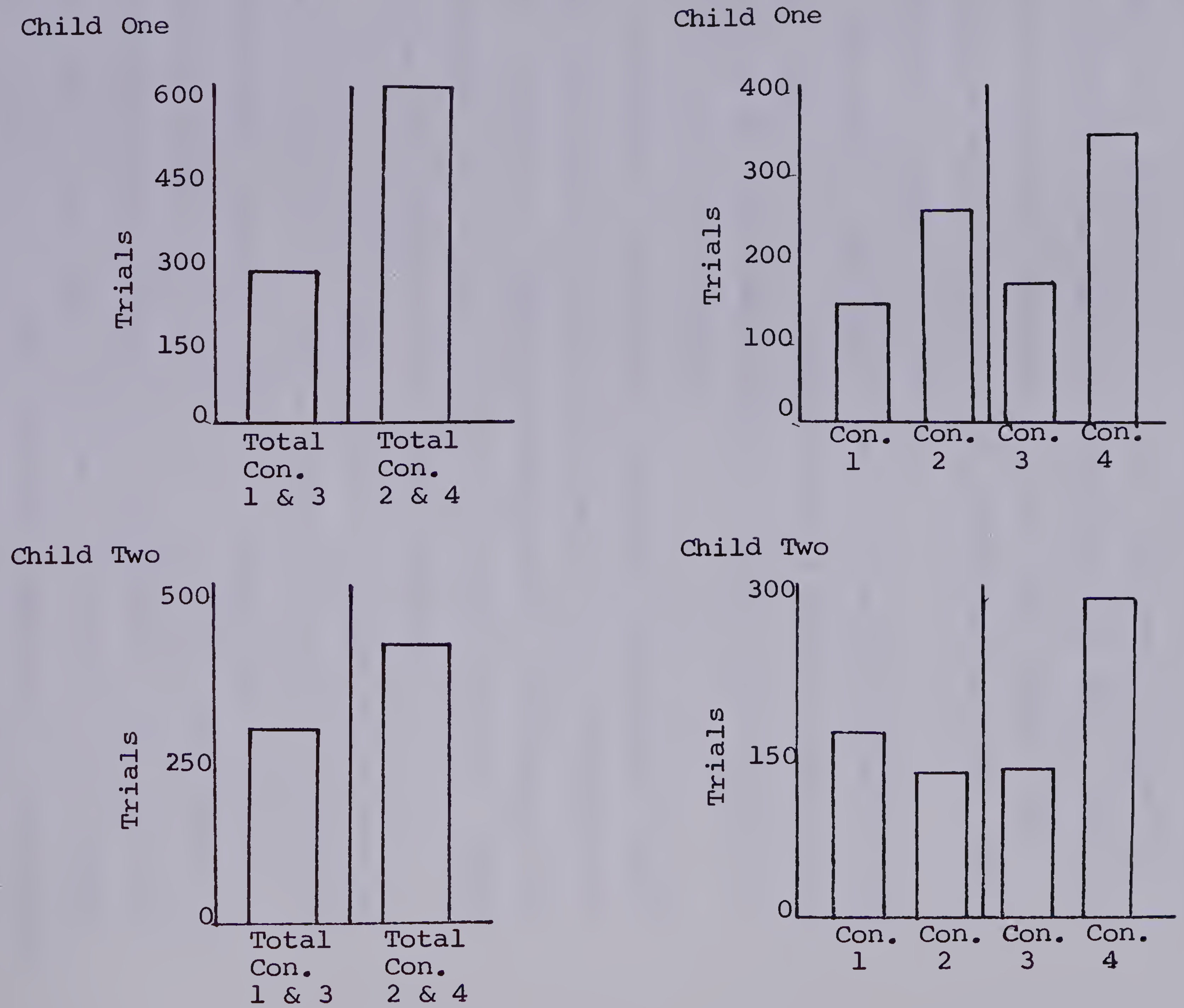
A further comparison is made for increasing the number of exemplars from two to four within each level of interstimulus variation. This is observed by comparing condition one (NE level 1, IER level 1) with condition two (NE level 2, IER Level 1) and comparing condition three (NE level 1, IER level 2) with condition four (NE level 1, IER level 2). These comparisons for both children also appear in Figure P4.

The results show a very similar trend to the results for hypothesis one. Child one showed an increase in TTC for condition two over one while child two showed the reverse. Both children demonstrated a greater TTC for condition four over condition three. These results were taken as support of hypothesis two at the overall comparison and at level two interstimulus variability. That is, as number of exemplars increased from two to four there was an overall increase in TTC. This effect was not observed across both children for increasing number of exemplars from two to four at level one interstimulus variation.

Generalization Results

As there are only two children in the pilot study their generalization results will be presented as a comparison between children on individual generalization tasks. These

Figure P4:


Child One
600
450
300
150
0
Trials
Total Con. 1 & 3
Total Con. 2 & 4
Child One
400
300
200
100
0
Trials
Con. 1
Con. 2
Con. 3
Con. 4
Child Two
500
250
0
Trials
Total Con. 1 & 3
Total Con. 2 & 4
Child Two
300
150
0
Trials
Con. 1
Con. 2
Con. 3
Con. 4

comparisons were stated in hypotheses four and six.

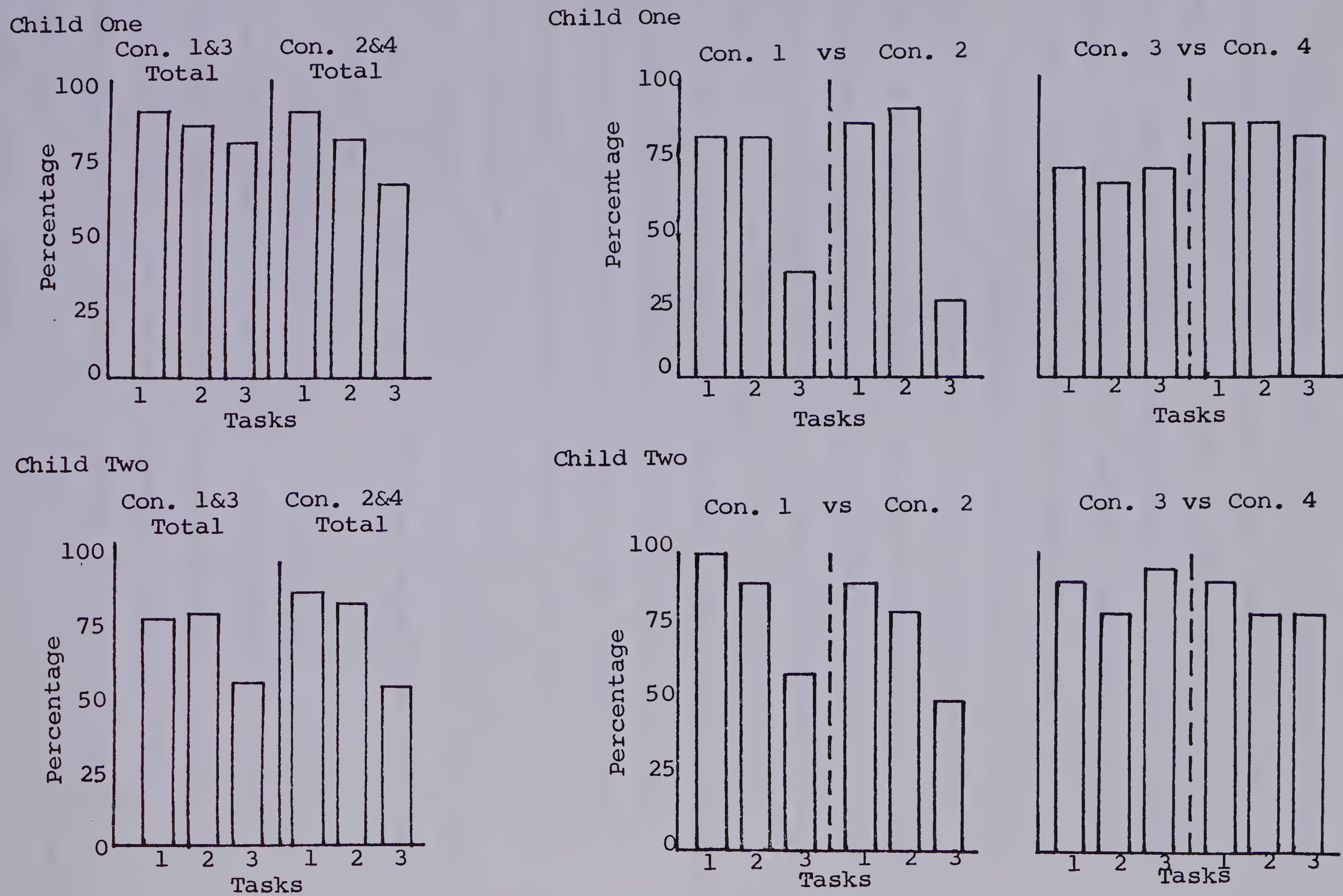
Hypothesis four stated that as the number of exemplars increased from two to four there would be a difference in the total generalization responses on each of the component tasks in the generalization test. The overall comparison for increasing the number of exemplars from two to four is made by comparing conditions one and three totals with conditions two and four at each generalization task. This comparison is seen in Figure P5.

The results show that there is no major difference between conditions one and three total and two and four total on any of the three generalization tasks. This effect was observed for both children.

A further comparison of increasing the number of exemplars from two to four at each level of interstimulus variation was made by comparing generalization task scores for conditions one and two and then for three and four. Figure P5 shows this comparison for both children.

The results show that there is no major difference between conditions one and two on any generalization task for either child. A similar effect is noted when comparing condition three and four for both children. These results were viewed as non-supportive of hypothesis four, that is, there was not any major difference on generalization tasks one, two and three as the number of exemplars increased from two to four, for either the overall comparison or at interstimulus level one and two.

Figure P5: Generalization scores to Tasks one two and three



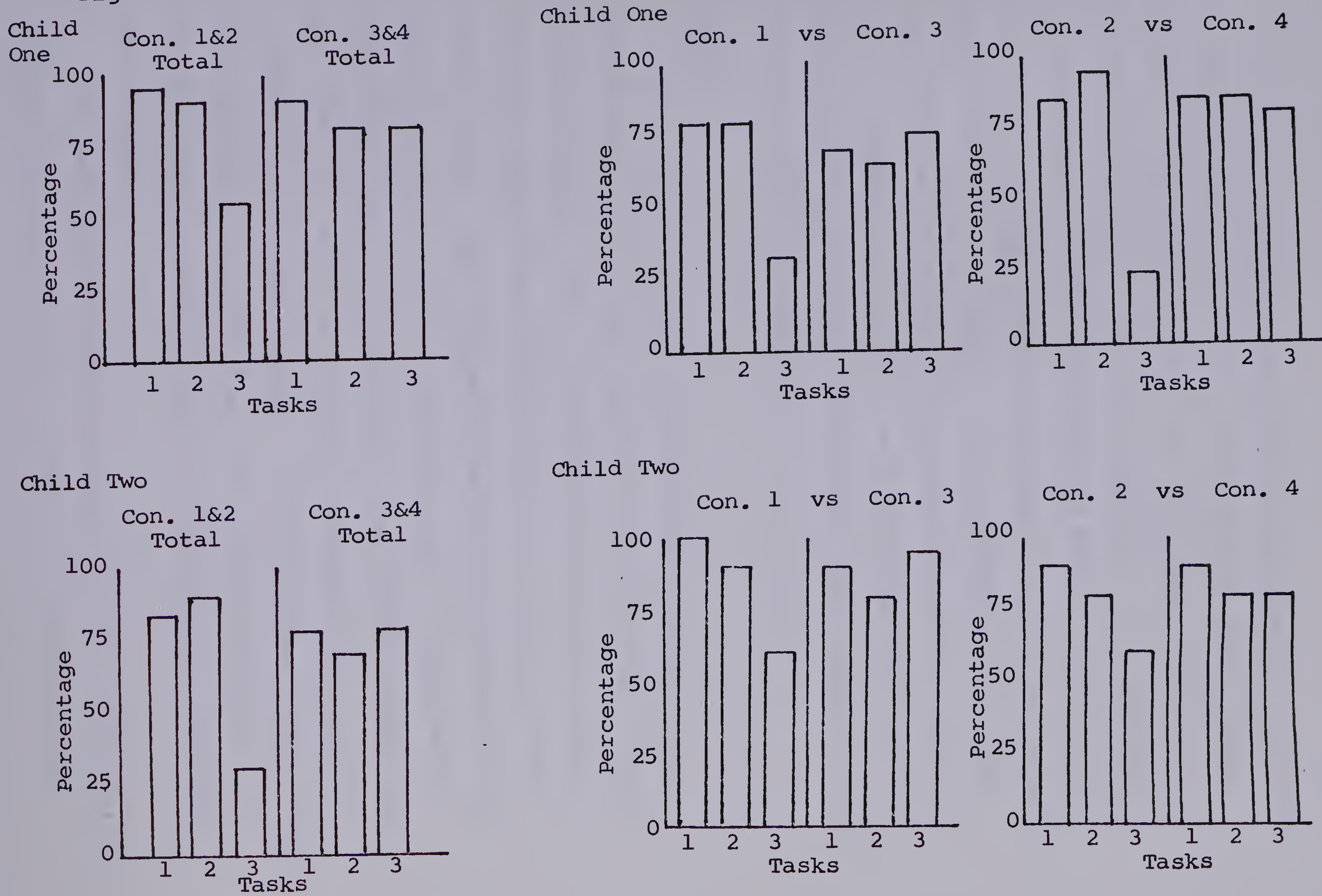
Hypothesis six stated that as interstimulus variation increased from one to three irrelevant dimensions there would be a difference in the total generalization responses to each task in the generalization test. The overall comparison for this hypothesis is made by comparing conditions one and two total with conditions three and four total.

The results of this comparison for child one and child two are presented in Figure P6. These results show that there is no major difference in conditions one and two total and three and four total at either generalization task one and two; while there is a major difference at generalization task three in favor of conditions three and four total.

Two further comparisons were made comparing the effects of increasing interstimulus variation from one to three irrelevant dimensions at each level of NE. This was made by comparing condition one (IER level 1, NE level 1) with c ondition three (IER level 2, NE level 1) and condition two (IER level 1, NE level 2) with condition four (IER level 2, NE, level 2) The results of these comparisons are shown in Figure P6.

As interstimulus variation increased from one to three irrelevant dimensions at level one NE (condition one with condition three) there was no major difference between condition one and three on generalization tasks one and two. There was a major difference, however for generalization task three in favor of condition three. This effect was observed for both children.

Figure P6: Generalization Scores to Tasks One, Two and Three



A similar trend in generalization results was observed for increasing interstimulus variation from one to three irrelevant dimensions at level two NE (condition two with condition four). The only difference was generalization task three in favor of condition four. These results were viewed as supportive of hypothesis six only on generalization task three. That is, as interstimulus variation increased from one to three irrelevant dimensions there was also an increase in the total generalization responses to generalization task three. This effect was observed at both levels of number of exemplars for both children.

Summary

The relation between acquisition and generalization for children one and two suggests a similar trend. Child one showed greater TTC for condition three over condition one while generalization was greater for condition three at task three. This suggests greater generalization from condition three, however, there was also an increase in acquisition time. Condition four also showed greater acquisition time over condition two with increased generalization on generalization task three. Condition three however, showed fewer TTC than condition four, while also resulting in a similar level of generalization to generalization task one, two and three.

These findings suggest that the maximum generalization from minimal training time is found in condition three.

This further suggests that diversity of exemplars (IER) during acquisition training rather than number of exemplars (NE) is the crucial dimension when programming for generalization. These results are consistent with the findings of the major study.

Subject two revealed a slightly different trend for acquisition and a similar trend for generalization results. Condition three took fewer TTC over condition one, while showing similar levels of generalization on tasks one, two and greater generalization on task three. In a similar manner to subject one, condition three also required fewer TTC over condition four with no major difference in generalization between the two conditions on tasks one, two and three.

As a result, condition three demonstrated the greatest savings in acquisition time with the greatest amount of generalization. The results for subject two are seen as a replication of the results for subject one, suggesting that diversity of exemplars rather than number of exemplars is the important dimension when programming for maximum generalization.

APPENDIX B

Description of stimulus materials comprising the teaching sets for each concept pair in the six instructional conditions

Concept Pair: Long-Short

Examples of Concept Instances in Four Instructional Conditions

Teaching Set	Materials Description	Concept Instance Long	Concept Instance Short
Condition One Teaching Set			
(1) 2 trains	color, form (same)	6 cars	2 cars
(2) 2 carrots	color, form (same)	6 cars	2 cars
Condition Two Teaching Set			
(1) 2 trains	same as condition One		
(2) 2 carrots	same as condition One		
(3) 2 shovels	color, form (same)	6 cars	2 cars
(4) 2 fish	color, form (same)	6 cars	2 cars
Condition Three Teaching Set			
(1) a. train	black; form, color different	2 x short	½ x long
b. dog	brown; form, color different	2 x short	½ x long
(2) a. shovel	blue; form, color different	2 x short	½ x long
b. fish	silver; form, color different	2 x short	½ x long
Condition Four Teaching Set			
(1) a. train	same as condition three		
b. dog	same as condition three		
(2) a. shovel	same as condition three		
b. fish	same as condition three		
(3) a. shoe	white/black; form, color different	2 x short	½ x long
b. snake	green; form, color different	2 x short	½ x long
(4) a. arrow	black; form, color different	2 x short	½ x long
b. hockey stick	yellow; form, color different	2 x short	½ x long

Concept Pair: Straight-Curved

Examples of Concept Instances in Four Instructional Conditions

Condition One Teaching Set	Materials Description	Concept Instance Straight	Concept Instance Curved
(1) 2 roads	color, form, length (same)	vertical 6"	curve 6"
(2) 2 elephants	color, form, length (same)	vertical 6"	curve 6"
Condition Two Teaching Set			
(1) 2 roads (2) elephants	same as condition one		
(3) 2 fish poles	color, form, length (same)	vertical 4"	curved 4"
(4) 2 flag poles	color, form, length (same)	vertical 4"	curved 4"
Condition Three Teaching Set			
(1) a. road b. fish pole	black yellow } length, same color, form different	vertical 4" vertical 4"	curved shape 4" curved 4"
(2) a. elephant b. flag pole	pink brown } length, same color, form different	vertical 4" vertical 4"	curved shape 4" curved shape 4"
Condition Four Teaching Set			
(1) a. road b. fish pole (2) a. elephant b. flag pole	same as condition three		
(3) a. train b. rope	black yellow } length, same color, form different	horizontal 4" horizontal 4"	curved shape 4" curved shape 4"
(4) a. giraffe b. pencil	green/yellow red/blue } length same color, form different	vertical 4" vertical 4"	curved shape 4" curved shape 4"

Concept Pair: Empty-Full

Examples of Concept Instances in Four Instructional Conditions

Condition One Teaching Set	Materials Description	Concept Instance Empty	Concept Instance Full
(1) 2 baskets	color, form, size, same	basket empty	2 kittens in
(2) 2 bird nests	color, form, size, same	bird nests empty	1 bird in
Condition Two Teaching Set			
(1) 2 baskets (2) 2 bird nests	same as condition one		
(3) 2 trucks	color, form, size, same	truck box empty	truck box full of samd
(4) 2 boats	color, form, size, same	boat empty	1 fisherman in
Condition Three Teaching Set			
(1) a. basket b. nest	brown green — form, color different	basket empty nest empty	2 kittens in 2 birds in
(2) a. box b. wagon	red black — form, color different	box empty wagon empty	1 toy in 2 kids in
Condition Four Teaching Set			
(1) a. basket b. nest (2) a. box b. wagon	same as condition three		
(3) a. egg carton b. playpen	blue/white red — form, color, different	no eggs no child	12 eggs in 1 child in
(4) a. truck b. sleigh	brown grey — form, color, different	no horse no kids	1 horse in 2 kids in

Concept Pair: In-Out

Examples of Concept Instances in Four Experimental Conditions

Condition One Teaching Set	Materials Description	Concept Instance In	Concept Instance Out
(1) a. bird nest with bird in	form, color, size same	in	
b. bird nest with bird out			out
(2) a. wheelbarrow with rabbit in	form, color, size same	in	
b. wheelbarrow with rabbit out			out
Condition Two Teaching Set			
(1) a. bird nest with bird in	same as condition one		
b. bird nest with bird out			
(2) a. wheelbarrow with rabbit in			
b. wheelbarrow with rabbit out			
(3) a. girl/pool with girl in	form, color, size same	in	
b. girl/pool with girl out			out
(4) a. car/boy with boy in	form, color, size same	in	
b. car/boy with boy out			out

Condition Three Teaching Set	Concept Instance		Concept Instance
(1) a. nest with bird in	in	a) nest with bird beside	out
b. wheelbarrow with rabbit beside	out	b) wheelbarrow with rabbit in	in
(2) a. basket with fruit in	in	c) basket with fruit beside	out
b. bottle with crayon beside	out	d) bottle with crayon in	in

Condition Four Teaching Set	Concept Instance		Concept Instance
(1), (2) same as condition three			
(3) a^3. swimming pool with girl in it	in	a^{31} swimming pool with girl beside it	out
b^3. car with boy beside it	out	b^{31} car with boy in it	in
(4) a^4. fish bowl with fish in it	in	a^{41} fish bowl with fish beside it	out
b^4. boat with boys beside it	out	b^{41} boat with boy in it	in

APPENDIX C

Arrangement of stimulus materials within teaching session for concept pairs in the four experimental conditions

Arrangement of stimulus materials within teaching session for concept pairs in the four experimental conditions

Each experimental condition differs on the number of exemplars representing a concept class and the number of irrelevant dimensions which vary between stimulus members of each exemplar. Within each condition the number of exemplars and irrelevant dimension between stimuli were represented equally within sixteen instructional trials.

Condition One

This condition had two pairs of exemplars and one irrelevan dimension which varied between stimuli pairs (position) on either a right or left position. The sixteen instructional trials consisted of eight trials for each exemplar and within each set of eight one member of the concept pair was on the right position four times and on the left position four times. These arrangements of position and number of exemplars were ordered from one to sixteen alternating position and concept pair exemplar.

Example concept pair: long-short

Left Position		Right Position	
Stimulus	Concept Instance	Stimulus	Concept Instance
1. train (6cars)	long	train (2 cars)	short
2. train (2 cars)	short	train (6 cars)	long
3. carrot (6")	long	carrot (2")	short
4. carrot (2")	short	carrot (6")	long
5. train (6 cars)	long	train (2 cars)	short
6. train (2 cars)	short	train (6 cars)	long
7. carrot (6")	long	carrot (2")	short

8. carrot (2")	short	carrot (6")	long
9. train (6 cars)	long	train (2 cars)	short
10. train (2 cars)	short	train (6 cars)	long
11. carrot (6")	long	carrot (2")	short
12. carrot (2")	short	carrot (6")	long
13. train (6 cars)	long	train (2 cars)	short
14. train (2 cars)	short	train (6 cars)	long
15. carrot (6")	long	carrot (2")	short
16. carrot (2")	short	carrot (6")	long

Using a list of random numbers the order of introduction for the sixteen instructional trials for any concept pair is as follows: 6, 7, 8, 10, 14, 13, 2, 11, 9, 1, 4, 3, 5, 15, 12, 16.

Given there was a total of thirty-two trials this randomized list of sixteen trials was presented twice. The following list shows how the stimulus items were arranged on each of the thirty-two trials and which stimulus item was requested on each trial. Over the thirty-two trials the following Gellerman series was used to decide which concept was requested on each trial: AABABBABABBABAABAABABBABABBABAAB. One member of the concept pair was assigned letter A while the other was assigned letter B.

		Left Position			Right Position		
Trial		Stimulus	Concept Instance	Request	Stimulus	Concept Instance	Request
1	17	train	short	17	train	long	1
1	18	carrot	long	2	carrot	short	18
3	19	carrot	short	3	carrot	long	19
4	20	train	short	20	train	long	4

5	21	train	short	5	train	long	21
6	22	train	long	22	train	short	6
7	23	train	short	23	train	long	7
8	24	carrot	long	24	carrot	short	8
9	25	train	long	9	train	short	25
10	26	train	long	26	train	short	10
11	27	carrot	short	11	carrot	long	27
12	28	carrot	long	12	carrot	short	28
13	29	train	long	29	train	short	13
14	30	carrot	long	14	carrot	short	30
15	31	carrot	short	31	carrot	long	15
16	32	carrot	short	16	train	long	32

Condition Two

This condition had four pairs of exemplars in the teaching set with one irrelevant dimension, position, varying between stimuli in each concept pair on either the right or left position. The possible arrangements for these four exemplars varied on position are presented in the following list.

Example concept pair: Straight-Curved

		Left Position		Right Position	
		Stimulus	Concept Instance	Stimulus	Concept Instance
A	1	Road	Straight	Road	Curved
	2	Road	Curved	Road	Straight
	3	Road	Straight	Road	Curved
	4	Road	Curved	Road	Straight

B	1	Elephant	Straight	Elephant	Curved
	2	Elephant	Curved	Elephant	Straight
	3	Elephant	Straight	Elephant	Curved
	4	Elephant	Curved	Elephant	Straight
C	1	Fish Pole	Straight	Fish Pole	Curved
	2	Fish Pole	Curved	Fish Pole	Straight
	3	Fish Pole	Straight	Fish Pole	Curved
	4	Fish Pole	Curved	Fish Pole	Straight
D	1	Flag Pole	Straight	Flag Pole	Curved
	2	Flag Pole	Curved	Flag Pole	Straight
	3	Flag Pole	Straight	Flag Pole	Curved
	4	Flag Pole	Curved	Flag Pole	Straight

A randomized list of the possible combinations was arranged by assigning each set of exemplars a letter (A, B, C, D). One trial was chosen from each of these sets according to the following series DABDCBADCBACDBCA. Within each set of four exemplars a rotating number series was established:

	Series			
A	1	2	3	4
B	2	3	4	1
C	3	4	1	2
D	4	1	2	3

Items were then arranged in a teaching list by choosing items first from the letter series and secondly from the number arrangement. This was done until sixteen arrangements were exhausted. Given there was a total of thirty-two trials this list of sixteen trials was presented twice. The following

list shows how the stimulus items were arranged on each of the thirty-two trials and which stimulus items were requested on each trial.

Trial		Left Position Stimulus	Left Position Concept Instance	Left Position Request	Right Position Stimulus	Right Position Concept Instance	Right Position Request
1	17	Flag Pole	Curved	1	Flag Pole	Straight	17
2	18	Road	Straight	2	Road	Curved	18
3	19	Elephant	Curved	19	Elephant	Straight	3
4	20	Flag Pole	Straight	4	Flag Pole	Curved	20
5	21	Fish Pole	Straight	21	Fish Pole	Curved	5
6	22	Elephant	Straight	22	Elephant	Curved	6
7	23	Road	Curved	7	Road	Straight	23
8	24	Flag Pole	Curved	24	Flag Pole	Straight	8
9	25	Fish Pole	Straight	9	Fish Pole	Curved	25
10	26	Elephant	Curved	26	Elephant	Straight	10
11	27	Road	Straight	27	Road	Curved	11
12	28	Fish Pole	Curved	12	Fish Pole	Straight	27
13	29	Flag Pole	Straight	29	Flag Pole	Curved	13
14	30	Elephant	Straight	14	Elephant	Curved	30
15	31	Fish Pole	Curved	15	Fish Pole	Straight	31
16	32	Road	Curved	32	Road	Straight	16

The request for concept instances on each trial was arrived at by the same procedure of assigning one concept member to A, the other to B, and following the Gellerman series listed in condition one.

Condition Three

This condition contained two pairs of exemplars in the teaching set with three irrelevant dimensions, position, color and form varying between stimuli in each concept pair. All

possible arrangements for these irrelevant dimensions varied between stimulus items are presented in the following list.

Example concept pair: Empty-full

		Left Position		Right Position	
		Stimulus	Concept Instance	Stimulus	Concept Instance
	1	Basket (red)	Full	Bird nest (blue)	Empty
	2	Basket (red)	Empty	Bird nest (blue)	Full
A	3	Basket (blue)	Full	Bird nest (red)	Empty
	4	Basket (blue	Empty	Bird nest (red)	Full
	1	Bird nest (blue)	Full	Basket (red)	Empty
B	2	Bird nest (blue)	Empty	Basket (red)	Full
	3	Bird nest (red)	Full	Basket (blue)	Empty
	4	Bird nest (red)	Empty	Basket (blue)	Full
	1	Box (brown)	Full	Wagon (green)	Empty
C	2	Box (brown)	Empty	Wagon (green)	Full
	3	Box (red)	Full	Wagon (blue)	Empty
	4	Box (redP	Empty	Wagon (blue)	Full
	1	Wagon (green)	Full	Box (brown)	Empty
D	2	Wagon (green)	Empty	Box (brown)	Full
	3	Wagon (blue)	Full	Box (red)	Empty
	4	Wagon (blue)	Empty	Box (red)	Full

Employing the same procedure described in condition two the following list of thirty-two trials shows the arrangement of stimulus materials and concept instances requested on each trial.

Trial		Left Position Stimulus	Concept Instance	Re-quest	Right Position Stimulus	Concept Instance	Re-quest
1	17	Wagon (blue)	Empty	17	Box (red)	Full	1
2	18	Basket (red)	Full	2	Nest (blue)	Empty	18

3	19	Nest (blue)	Empty	3	Basket (red)	Full	19
4	20	Wagon (green)	Full	20	Box (brown)	Empty	4
5	21	Box (red)	Full	5	Wagon (blue)	Empty	21
6	22	Nest (green)	Full	22	Basket (blue)	Empty	6
7	23	Basket (red)	Empty	23	Nest (blue)	Full	7
8	24	Wagon (green)	Empty	24	Box (brown)	Full	8
9	25	Box (red)	Empty	9	Wagon (blue)	Full	25
10	26	Nest (green)	Empty	26	Basket (blue)	Full	10
11	27	Basket (blue)	Full	11	Nest (green)	Empty	27
12	28	Box (brown)	Full	12	Wagon (green)	Empty	28
13	29	Wagon (blue)	Full	29	Box (red)	Empty	13
14	30	Nest (blue)	Full	14	Basket (red)	Empty	30
15	31	Box (brown)	Empty	31	Wagon (green)	Full	15
16	32	Basket (blue)	Empty	16	Nest (green)	Full	32

Full = A, Empty = B

Condition Four

This condition contained four pairs of exemplars in the teaching set and three irrelevant dimensions, position, color and form, varying between stimuli in each concept pair. All possible arrangements for these irrelevant dimensions varied between stimulus items are presented in the following list. Given that there are thirty-two possible arrangements, two separate series had to be arranged: one is labelled A, B, C, D, while the second is labelled A^1, B^1, C^1, D^1.

Example concept pair: In-Out

		Left Position		Right Position	
		Stimulus	Concept Instance	Stimulus	Concept Instance
A	1	nest (brown,yellow)	in	wheelbarrow (red)	out
	2	nest (brown,yeallow)	out	wheelbarrow (red)	in
	3	nest (black,orange)	in	wheelbarrow (purple)	out
	4	nest (black,orange)	out	wheelbarrow (purple)	in

	1	wheelbarrow (red)	in	nest (brown,yellow)	out
	2	wheelbarrow (red)	out	nest (brown,yellow)	in
A^1	3	wheelbarrow (purple)	in	nest (black,orange)	out
	4	wheelbarrow (purple)	out	nest (black,orange)	in
	1	girl/pool (blue)	in	jar/balls (red)	out
B	2	girl/pool (blue)	out	jar/balls (red)	in
	3	girl/pool (red)	in	jar/balls (green)	out
	4	girl/pool (red)	out	jar/balls (green)	in
	1	jar/balls (red)	in	girl/pool (blue)	out
B^1	2	jar/balls (red)	out	girl/pool (blue)	in
	3	jar/balls (green)	in	girl/pool (red)	out
	4	jar/balls (green)	out	girl/pool (red)	in
	1	shoe/pencil (blue)	in	dog/tub (green)	out
	2	shoe/pencil (blue)	out	dog/tub (green)	in
C	3	shoe/pencil (red)	in	dog/tub (blue)	out
	4	shoe/pencil (red)	out	dog/tub (blue)	in
	1	dog/tub (green)	in	shoe/pencil (blue)	out
	2	dog/tub (green)	out	shoe/pencil (blue)	in
C^1	3	dog/tub (blue)	in	shoe/pencil (red)	out
	4	dog/tub (blue)	out	shoe/pencil (red)	in
	1	car/boy (orange)	in	bottle (blue)	out
	2	car/boy (orange)	out	Bottle (blue)	in
D	3	car/boy (green)	in	bottle (orange)	out
	4	car/boy (green)	out	bottle (orange)	in
	1	bottle (blue)	in	car/boy (orange)	out
	2	bottle (blue)	out	Car/boy (orange)	in
D^1	3	bottle (blue)	in	car/boy (green)	out
	4	bottle (orange)	out	car/boy (green)	in

Employing the procedure described in condition two for each series (A B C D and A^1 B^1 C^1 D^1) two separate lists were

arranged. These lists were labelled series A and series B and were administered alternately over instructional sessions. An example of Series A is shown in the following list.

Series A

		Left Position			Right Position		
Trial		Stimulus	Concept Instance	Re-quest	Stimulus	Concept Instance	Re-quest
1	17	Car (green)	out	17	bottle (orange)	in	1
2	18	nest (brown)	in	2	wheelbarrow (red)	out	18
3	19	girl/pool (blue)	out	3	jar/balls (red)	in	19
4	20	car/boy(orange)	in	20	bottle (blue)	out	4
5	21	shoe/pencil(red)	in	5	dog/tub (blue)	out	21
6	22	girl/pool (red)	in	22	jar/balls (green)	out	6
7	23	nest (brown)	out	23	wheelbarrow (red)	in	7
8	24	car/boy(orange)	out	24	bottle (blue)	in	8
9	25	shoe/pencil (red)	out	9	dog/tub (blue)	in	25
10	26	girl/pool (red)	out	26	jar/balls(green)	in	10
11	27	nest (black)	in	11	wheelbarrow(purple)	out	27
12	28	shoe/pencil(blue)	in	12	dog/tub(green)	out	28
13	29	car/boy (green)	in	29	bottle (orange)	out	13
14	30	girl/pool(Blue)	in	14	jar/balls(red)	out	30
15	31	shoe/pencil(blue)	out	31	dog/tub (green)	in	15
16	32	nest (black)	out	16	wheelbarrow (purple)	in	32

APPENDIX D

Description of stimulus sets for each concept pair on the generalization test (condition 1, 2, 3)

APPENDIX D

Concept Pair: In-out

Condition One	Materials Concept Pairs	Concept Instance	Attributes Form	Attributes Color	Attributes Size
Set 1.	(1) cup with pencil in	in	same	same	same
	(2) cup with pencil beside	out			
Set 2.	(1) box with ball in	in	"	"	"
	(2) box with ball beside	out			
Set 3.	(1) hat with dog in	in	"	"	"
	(2) hat with dog beside	out			
Set 4.	(1) playpen with child in	in	"	"	"
	(2) playpen with child beside	out			
Set 5.	(1) pail with shovel in	in	"	"	"
	(2) pail with shovel beside	out			
Condition Two					
Set 1.	(1) bed with child in	in	"	child pink	"
	(2) bed with child beside	out	"	child brown	"
Set 2.	(1) flowerpot with flower in	in	"	flower yellow	"
	(2) flowerpot with flower beside	out	"	flower red	"
Set 3.	(1) fruit bowl with fruit in	in	"	red apples	"
	(2) fruit bowl with fruit beside	out	"	green apples	"
Set 4.	(1) wagon with blocks in	in	"	blocks blue	"
	(2) wagon with blocks beside	out	"	blocks red	"
Set 5.	(1) doghouse, dog in	in	"	dog, white/ black	"
	(2) doghouse, dog beside	out	"	dog, brown	"

Condition Three	Materials Concept Pairs	Concept Instance	Attributes Form	Color	Size
Set 1.	(1) hat with pencil in	in	different	different	same
	(2) egg cup with egg besdie	out	"	"	"
Set 2.	(1) glass with snake in	in	"	"	"
	(2) icecream cone with ice-cream beside	out			
Set 3.	(1) washtub with pigs in	in	"	"	"
	(2) waste paper basket with paper beside	out			
Set 4.	(1) purse with shoe in	in	"	"	"
	(2) bath tub with child beside	out			
Set 5.	(1) bed with child in	in	"	"	"
	(2) box with pots beside	out			

Concept Pair: Empty-full

Condition One	Materials Concept Pairs	Attributes	Concept Instance empty		Concept Instance full
Set (1)	2 bottles	1. color	-	same	-
		2. size/form	-	same	-
		3. contents	none	different	¾ full
Set (2)	2 fish tanks	1. color	-	same	-
		2. size/form	-	same	-
		3. contents	none	different	3 fish
Set (3)	2 bowls	1. color	-	same	-
		2. size/form	-	same	-
		3. contents	none	different	apples (fruit)
Set (4)	2 clear jars	1. color	-	same	-
		2. size/form	-	same	-
		3. contents	none	different	¾ full candies
Set (5)	2 benches	1. color	-	same	-
		2. size/form	-	same	-
		3. contents	none	different	2 people sitting
Condition Two					
Set (1)	2 egg cups	1. color	red	different	blue
		2. size/form	-	same	-
		3. contents	none	different	one egg
Set (2)	2 flower pots	1. color	green	different	yellow
		2. size/form	-	same	-
		3. contents	none	different	1 flower
Set (3)	2 open hands	1. color	pink	different	brown
		2. size/form	-	same	-
		3. contents	none	different	2 blocks

Set (4)	2 shopping carts	1. color	black	different	red
		2. size/form	-	same	-
		3. contents	none	different	¾ full groceries
Set (5)	2 cars	1. color	green	different	yellow
		2. size/form	-	same	-
		3. contents	none	different	one person

Condition Three

		Concept Instance	Attributes: Color	Attributes: Relative size
Set (1)	a. swimming pool with children	full	different	same
	b. sandbox empty	empty		
Set (2)	a. wagon with balls and blocks	full	different	same
	b. glass empty	empty		
Set (3)	a. bed with two children in	full	different	same
	b. bird cage	empty		
Set (4)	a. box	empty	different	same
	b. square fence with 2 cows	full		
Set (5)	a. bath tub with child	full	different	same
	b. wash tub empty	empty		

Concept Pair: Straight-curved

Condition One	Materials Concept Pairs	Attributes	Concept Instance straight		Concept Instance curved
Set (1)	2 pieces of string	1. color	-	same	-
		2. length/form	-	same	-
		3. shape	straight	different	curved in c shape
Set (2)	2 trees	1. color	-	same	-
		2. length/form	-	same	-
		3. shape	straight	different	curved in c shape
Set (3)	2 snakes	1. color	-	same	-
		2. length/form	-	same	-
		3. shape	straight	different	curved in c shape
Set (4)	2 belts	1. color	-	same	-
		2. length/form	-	same	-
		3. shape	straight	different	curved in c shape
Set (5)	2 fingers	1. color	-	same	-
		2. length/form	-	same	-
		3. shape	straight	different	approx. c shape
Condition Two					
Set (1)	2 candles	1. color	red	different	yellow
		2. length/form	-	same	-
		3. shape	straight	different	curved in c shape
Set (2)	2 fish	1. color	multicolored	different	multicolored
		2. length/form	-	same	-
		3. shape	straight	different	curved in c shape
Set (3)	2 arms	1. color	pink	different	black
		2. length/form	-	same	-
		3. shape	straight	different	approx. c shape

Set (4)	2 scarves	1. color	orange/blue	different	brown/grey
		2. length/form	-	same	-
		3. shape	straight	different	curved in c shape
Set (5)	2 boys	1. color	clothes diff. colors	different	clothes diff. colors
		2. length/form	-	same	-
		3. shape	straight	different	curved in c shape

Condition Three

		Concept Instances	Attributes: color	Attributes: length
Set (1)	a. flag pole	straight	different	same
	b. letter C	curved		
Set (2)	a. ruler	straight	different	same
	b. horse shoe	curved		
Set (3)	a. coat rack	straight	different	same
	b. archway	curved		
Set (4)	a. fishing rod	straight	different	same
	b. cresent moon	curved		
Set (5)	a. flower	straight	different	same
	b. banana	curved		

Concept Pair: Long-short

Condition One	Materials Concept Pair	long Concept Instance	short Concept Instance
Set (1)	2 pencils	1. color, yellow	1. same as long
		2. shape, thin pointed	2. same as long
		3. length, 2 x short	3. length, ½ x long
Set (2)	2 cars	1. multicolored	1. same as long
		2. shape	2. same as long
		3. length, 2 x short	3. length, ½ x long
Set (3)	2 ladders	1. color, red	1. same as long
		2. shape, thin	2. same as long
		3. length, 2 x short	3. ½ x long
Set (4)	2 flagpoles	1. color, multi	1. same as long
		2. shape	2. same as long
		3. length, 2 x short	3. length, ½ x long
Set (5)	2 spoons	1. color, grey	1. same as long
		2. teaspoon	2. same as long
		3. length, 2 x short	3. length, ½ x long
Condition Two			
Set (1)	2 candles	1. yellow	1. red
		2. shape (½" x 4" c flame)	2. same as long
		3. length, 2 x short	3. length, ½ x long
Set (2)	2 rulers	1. red	1. blue
		2. shape (½" x 4", vert.)	2. same as long
		3. 2 x short	3. ½ x long
Set (3)	2 forks	1. black	1. green
		2. table fork	2. same as long
		3. 2 x short	3. ½ x long

Set (4)	2 ropes	1. grey 2. long, thin 3. 2 x short	1. yellow 2. same as long 3. ½ x long
Set (5)	2 tables	1. brown 2. sitting table 3. 2 x short	1. green 2. same as long 3. ½ x long

Condition Three		Concept Instance	Attributes: color	Attributes: length
Set (1)	a. boat b. bus	long short	different	½ x bus 2 x boat
Set (2)	a. ladder b. baseball bat	short long	different	½ x bat 2 x ladder
Set (3)	a. knife b. spoon	long short	different	2 x spoon ½ x knife
Set (4)	a. comb b. wagon	short long	different	½ x wagon 2 x comb
Set (5)	a. toboggan b. sleigh	short long	different	½ x sleigh 2 x toboggan